BMJ Open

Severe COVID-19 is associated with increased incidence of long-term respiratory, cardiovascular, and mental health conditions

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-056284
Article Type:	Original research
Date Submitted by the Author:	30-Aug-2021
Complete List of Authors:	Jovanoski, Nick; F Hoffmann-La Roche Ltd Chen, Xin; Genentech Inc Becker, Ursula; F Hoffmann-La Roche Ltd Zalocusky, Kelly; Genentech Inc Chawla, Devika; Genentech Inc Tsai, Larry; Genentech Inc Borm, Michelle; F Hoffmann-La Roche Ltd Neighbors, Margaret; Genentech Inc Yau, Vincent; Genentech Inc
Keywords:	COVID-19, Epidemiology < INFECTIOUS DISEASES, PUBLIC HEALTH, Respiratory infections < THORACIC MEDICINE

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Severe COVID-19 is associated with increased incidence of long-term respiratory, cardiovascular, and mental health conditions

Nick Jovanoski[1], Xin Chen[2], Ursula Becker[1], Kelly Zalocusky[2], Devika Chawla[2], Larry Tsai[2], Michelle Borm[3], Margaret Neighbors[2], Vincent Yau[2]

1. F. Hoffmann-La Roche Ltd, Basel, Switzerland

2. Genentech, Inc., South San Francisco, CA, USA

3. Roche Nederland BV, Woerden, Netherlands.

Author for correspondence: Vincent Yau, Ph.D. Genentech, Inc., South San Francisco, CA, USA. yau.vincent@gene.com

Target journal: BMJ Open

Word count: 3548/4000

Figures / Tables: 4 / 2 (plus 1 / 6 supplementary)

References: 28

Abstract (300/300 words)

Objective: To identify potential risk factors for the occurrence of adverse long-term outcomes (LTOs) associated with coronavirus disease 2019 (COVID-19), using a large electronic health record (EHR) database.

Design: Retrospective cohort study. Patients with COVID-19 were assigned into sub-cohorts according to the most intensive treatment setting experienced. Newly diagnosed conditions were classified as respiratory, cardiovascular, or mental health LTOs at either >30–≤90 days or >90–≤180 days after COVID-19 diagnosis or hospital discharge. Multivariate regression analysis was performed to identify any effect of disease severity on LTO incidence.

Setting: Optum® de-identified COVID-19 EHR dataset drawn from hospitals and clinics across the United States.

Participants: Individuals diagnosed with COVID-19 (N=57,748) between February 20, 2020 and December 31, 2020.

Main outcomes: Incidence of new clinical conditions after COVID-19 diagnosis or hospital discharge and the potential effect of disease severity on their risk of occurrence.

Results: Patients were assigned into one of six sub-cohorts: outpatient (n=22,788), emergency room (ER) with same-day COVID-19 diagnosis (n=11,633), ER with COVID-19 diagnosis ≤21 days before ER visit (n=2,877), hospitalization without intensive care unit (ICU; n=16,653), ICU without ventilation (n=1,837), and ICU with ventilation (n=1,960). Respiratory LTOs were more common than cardiovascular or mental health LTOs across sub-cohorts, and LTO incidence was higher in hospitalized versus non-hospitalized sub-cohorts. Patients with the most severe disease (ICU with ventilation sub-cohort) were at increased risk of respiratory (risk ratio [RR] 1.86, 95% confidence interval [CI] 1.56, 2.21), cardiovascular (RR 2.65, 95% CI 1.49, 4.43), and mental health outcomes (RR 1.52, 95% CI 1.20, 1.91) up to six months after hospital discharge compared with outpatients.

Conclusions: Patients with severe COVID-19 had increased risk of new clinical conditions being diagnosed up to six months after hospital discharge. Strategies to prevent disease progression may reduce the risk of LTOs in patients with COVID-19.

Strengths and limitations of this study

- This study used a large electronic health record database containing a rich source of patient-level medical and administrative records from hospitals, emergency departments, and outpatient centers across the United States.
- Multivariate logistic regression analysis was used to adjust for measured confounders and assess the effect of increasing COVID-19 severity (proxied by treatment setting) on the risk of new clinical conditions being diagnosed up to six months after COVID-19 diagnosis or hospital discharge.
- A sensitivity analysis assessing the effect of increasing COVID-19 severity on the risk of a new cancer diagnosis served as a negative control.
- The main limitation of the study is that we use treatment setting as a proxy for COVID-19 severity, and therefore it is difficult to tease out effects specific to the treatment setting (e.g., invasive ventilation) from the underlying COVID-19 severity
- Additional limitations include missing information on smoking status, the lack of a COVID-19-negative control group, the possibility of missing data, being restricted to examining conditions captured by ICD-10 codes, the lack of information on COVID-19 treatments received, and the lack of laboratory values or other biomarkers to better characterize disease.

Background

The coronavirus disease 2019 (COVID-19) pandemic caused by the novel severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) has imposed an immense burden of morbidity and mortality worldwide.[1] Although the majority of patients experience mild or moderate symptoms that resolve within a few weeks of initial infection, increasing evidence suggests that a subset of patients continue to display symptoms beyond four weeks after infection.[2 3] These symptoms are wide ranging and often extend beyond the typical initial symptoms of COVID-19 to include respiratory (e.g., dyspnea, decreased exercise capacity), cardiovascular (e.g., heart palpitations, chest pain), and mental health (e.g., confusion, disorientation) disorders.[4 5] Notably, such outcomes have been observed even in patients with mild acute COVID-19 symptoms.[6] These prolonged symptoms have collectively been referred to by several names including post-acute COVID-19 (PAC), post-COVID-19 syndrome (PCS), post-acute sequelae of SARS-CoV-2 infection (PASC), and possibly more commonly 'long COVID'.[7 8] However, due to the overlapping and non-specific range of symptoms experienced, the medical community has not yet converged on precise definitions, and it is possible that distinct subsets of long COVID patients exist. It has also been suggested that long COVID can be further sub-divided into subacute COVID-19 (4–12 weeks after initial onset of COVID-19 symptoms) and post-COVID-19 syndrome (beyond 12 weeks).[4 9] The underlying pathogenic mechanisms of long COVID are not well understood, but multiple causes have been proposed, including immune dysregulation and viral persistence.[10] Additionally, in patients with severe disease requiring treatment in the intensive care unit (ICU), non-specific secondary effects cannot be ruled out, similar to those observed in 'post-intensive care syndrome'.[11]

High-quality clinical data on respiratory, cardiovascular, and neurologic sequelae of SARS-CoV-2 infection are beginning to emerge,[12-14] and several observational studies and patient registries have been established to better understand the long-term outcomes (LTOs)

of COVID-19.[15 16] However, little is known about the potential baseline factors that may predict the development of long COVID.

Retrospective cohort studies using electronic health records (EHRs) are uniquely positioned due to their size and convenience to provide insights into factors underlying long COVID development and the range of long COVID conditions that exist. The Optum® de-identified COVID-19 EHR dataset contains patient-level medical and administrative records from hospitals, emergency departments, outpatient centers, and laboratories across the United States (US). This dataset has previously been utilized to describe key epidemiological features of a large cohort of hospitalized patients with COVID-19[17] and to develop a prognostic model of in-hospital mortality.[18]

The current study utilized the Optum® de-identified COVID-19 EHR dataset to better understand the types of LTOs encountered by patients with long COVID, to define the factors that predict their diagnosis, and to understand the role COVID-19 severity plays in the manifestation of these outcomes.

Methods

Database

Individuals with COVID-19 diagnosed between February 20, 2020 and December 31, 2020 were extracted from the Optum® de-identified COVID-19 EHR dataset (569,149 individuals from 3,832,315 in the entire dataset). This dataset contains patient-level medical and administrative records from hospitals, emergency departments, outpatient centers, and laboratories across the US. All data were de-identified according to the Health Insurance Portability and Accountability Act Expert Method and managed according to Optum® customer data use agreements. The COVID-19 EHR dataset comprises clinical information sourced from hospital networks that provided data meeting Optum®'s internal data quality criteria. Data cleaning methods used were as described previously.[17]

Patients and study design

Eligible patients (overall COVID-19 cohort) had ≥1 of the following: a COVID-19 diagnosis code (U07.1, U07.2), a positive diagnostic test for SARS-CoV-2 infection (e.g., molecular or antigen test), or a B97.29 diagnosis code (other coronavirus as the cause of diseases classified elsewhere) without a negative SARS-CoV-2 molecular test within 14 days. The index date was defined as the date of COVID-19 diagnosis or COVID-19-related hospitalization (as defined below), whichever occurred first. The baseline period was defined as the 12 months prior to the index date, and a minimum of 180 days follow-up was required for all patients. The overall study design is shown in **Figure 1**.

Eligible patients were assigned into the following six sub-cohorts according to treatment setting: **1. Outpatient**, patients with a COVID-19 diagnosis and no record of hospitalization or an emergency room (ER) visit within 21 days of diagnosis; **2. ER on diagnosis**, COVID-19 diagnosis on the same day as ER visit; **3. ER**, COVID-19 diagnosis prior to ER visit, i.e., patients with an ER visit within 21 days after COVID-19 diagnosis (excluding diagnosis date); **4. Hospitalization without ICU**, patients hospitalized with no

record of ICU admission; **5. Hospitalized with ICU but no ventilation**, patients hospitalized with record of ICU admission but no record of ventilator or extracorporeal molecular oxygen (ECMO) use during ICU stay; **6. Hospitalized with ICU and ventilation**, patients hospitalized with record of ICU admission and ventilator or ECMO use during ICU stay.

Hospitalization was defined as an inpatient or ER overnight visit with an initial COVID-19 diagnosis made during hospitalization and within seven days of admission, or an inpatient or ER overnight visit within 21 days of the initial COVID-19 diagnosis, where the hospital had a record of this diagnosis. Contiguous ER and inpatient visits with a gap of up to one day were considered a single hospitalization. If a patient had multiple eligible hospitalizations, only data from the first hospitalization were considered, as described previously.[17]

Modeling and statistical analysis

LTOs occurring >30–≤180 days after hospital discharge or COVID-19 diagnosis were categorized into one of two time windows (>30–≤90 days or >90–≤180 days) and were further classified as respiratory, cardiovascular, or mental health conditions (**Supplemental Table 1**).[19] Multivariate logistic regression analyses were performed to determine the effect of disease severity (proxied by treatment setting) on the three LTO classifications. Covariates were intended to encompass the main known risk factors for developing severe COVID-19,[20] and included demographic information (i.e., age, gender, race, ethnicity, diagnosis month, insurance type, obesity status) and baseline health conditions (i.e., those included in the Charlson Comorbidity Index [CCI] (**Supplemental Table 2**). CCI was treated as a numeric variable, while all other variables were treated as categorical. Age was binned into <18 years, 18–29, 30–39, 40–49, 50–65, 65–74, 75–84, ≥85 years. Date of diagnosis was also binned into months in 2020 (pre-April, April, May, June, July). Patients were excluded from the regression model examining a specific LTO category if they had a diagnosis in that category in the 12 months prior to the index date (for example, if a patient

had an asthma diagnosis 12 months prior to the index date, they would be excluded from the model for respiratory LTOs).

All statistical analysis was performed using R 3.6.3.[21] Using the sjstats package, regression was performed using the function 'glm' and the risk ratio (RR) was calculated by converting the odds ratio (OR) using the function 'OR to RR'.[22] Increased risk of diagnosis of a health condition was implied when the RR and both the low and high 95% confidence interval limits (CI) were >1, and decreased risk was implied when the RR and low and high 95% CIs were <1.

Sensitivity analysis

A sensitivity analysis was performed to investigate the potential effect of disease severity (proxied by treatment setting) on risk of a new cancer diagnosis, to serve as a negative control. The same set of covariates was used as per the main analysis, but cancer diagnosis was the only LTO examined. Currently, no evidence exists to suggest that COVID-19 severity increases the risk of a new cancer diagnosis. Thus, an effect here may indicate that the effects from the main analysis may be driven by other differences between patients across treatment settings.

Patient and Public Involvement

No patient involved.

Results

Patient population

In total, 57,748 patients were eligible for the overall COVID-19 cohort. **Table 1** presents descriptive statistics of the patients by sub-cohort. Mean age tended to be higher in patients in hospitalized sub-cohorts (53.2–57.7 years) than in those in non-hospitalized sub-cohorts (41.0–46.8 years). Overall, 53.3% of patients were female. Across all patients, 50.3% were Caucasian, 22.8% were African American, 3.2% were Asian, and the remaining 23.6% were missing information on race. Additionally, 67.5% were of non-Hispanic ethnicity, while data on ethnicity was missing for 11.8% of patients. Overall, 19% of patients were obese and the mean weighted CCI score was 1.20. Information on smoking status was missing for 93.1% of patients (**Table 1**). Full details of demographics and baseline characteristics are provided in **Supplemental Table 3**.

The proportions of patients with incipient respiratory, cardiovascular, and / or mental health conditions that were diagnosed either >30–≤90 days or >90–≤180 days after COVID-19 diagnosis or hospital discharge are provided in **Table 2**. The proportions of patients with new LTOs were generally higher in the sub-cohorts with more severe disease (i.e., the ER sub-cohort and all hospitalized sub-cohorts) compared with the outpatient sub-cohort. In addition, the proportion of patients with respiratory LTOs was higher than the proportions with cardiovascular or mental health LTOs. New respiratory LTOs were diagnosed more frequently during the earlier time window across sub-cohorts, except in the outpatient sub-cohort where the proportion of patients diagnosed was the same in both time windows (both 8.1%; **Table 2**). No clear temporal trends were noted for diagnosis of cardiovascular or mental health LTOs, with similar proportions of patients with new cardiovascular and mental health LTOs observed in the >30–≤90- and >90–≤180-day windows for each sub-cohort (**Table 2**). The proportions of patients with LTOs in more than one category (i.e., 'respiratory and cardiovascular', 'respiratory and mental health', 'mental health and cardiovascular', or 'respiratory, cardiovascular, and mental health') were lower than the proportions of patients

with LTOs in a single category, indicating that a diagnosis in one category did not necessarily lead to a diagnosis in another.

Regarding individual conditions, the prevalence of newly diagnosed pneumonia, dyspnea, and respiratory failure in the >90–≤180-day window closely followed the pattern of initial COVID-19 severity, with most cases being diagnosed in the 'ICU with ventilation' sub-cohort (**Supplemental Table 4**). Similarly, although encephalopathy, confusion or disorientation, cardiac arrhythmia, and myocardial infarction were less common, the prevalence of these conditions also increased with increasing COVID-19 severity. Full details of conditions that were diagnosed in the >30–≤90- and >90–≤180-day windows following COVID-19 diagnosis or hospital discharge are provided in **Supplemental Table 4**.

Modeling

The most striking potential covariate associated with increased risk of newly diagnosed respiratory conditions at >30–≤90 days and >90–≤180 days post COVID-19 diagnosis or hospital discharge was increasing severity of illness according to increasing hospitalization severity, utilizing the outpatient sub-cohort as the reference group (**Figure 2** and **Supplemental Table 5**). ICU with ventilation was associated with increased risk of a novel respiratory condition diagnosis compared with the outpatient sub-cohort at >30–≤90 days (RR 2.64, 95% CI 2.27, 3.04) and >90–≤180 days post COVID-19 diagnosis or hospital discharge (RR 1.86, 95% CI 1.55, 2.21); in addition, ICU without ventilation was associated with increased risk during the >30–≤90-day time window (RR 1.69, 95% CI 1.39, 2.03), while ER was associated with increased risk at both >30–≤90 days (RR 1.39, 95% CI 1.17, 1.65) and >90–≤180 days (RR 1.33, 95% CI 1.10, 1.58) post COVID-19 diagnosis or hospital discharge. By contrast, patients with an ER visit on the COVID-19 diagnosis date were less likely than those in the outpatient sub-cohort to be diagnosed with a new respiratory condition at >30–≤90 days (RR 0.64, 95% CI 0.56, 0.74) and 90–180 days post COVID-19 diagnosis or hospital discharge (RR 0.56, 95% CI 0.48, 0.65). Additional covariates associated with increased risk of new respiratory conditions were older patient age and

obesity. A COVID-19 diagnosis during or prior to April 2020 exhibited a non-significant trend towards increased risk of new respiratory condition occurrence compared with later diagnosis, which may reflect changes in treatment algorithms over time. Full results are presented in **Supplemental Table 5**.

Increasing hospitalization severity was also found to be associated with increased risk of a new cardiovascular condition occurring post COVID-19 diagnosis or hospital discharge (**Figure 3** and **Supplemental Table 5**). Notably, ICU with ventilation was associated with increased risk of the occurrence of novel cardiovascular conditions compared with the outpatient sub-cohort at >30–≤90 days (RR 3.16, 95% CI 1.83, 5.18) and >90–≤180 days post COVID-19 diagnosis or hospital discharge (RR 2.65, 95% CI 1.49, 4.43), while ICU without ventilation was associated with increased risk during the >90–≤180-day time window (RR 2.41, 95% CI 1.25, 4.23). Similar to the findings regarding respiratory conditions, patients with an ER visit on the COVID-19 diagnosis date were less likely than outpatients to be diagnosed with novel cardiovascular conditions in both the >30–≤90-day (RR 0.45, 95% CI 0.27, 0.71) and >90–≤180-day windows (RR 0.59, 95% CI 0.38, 0.89). Additional covariates associated with an increased risk of new cardiovascular conditions occurring included older patient age and non-Hispanic ethnicity. Full results are presented in **Supplemental Table 5**.

The risk of a new mental health condition occurring post COVID-19 diagnosis or hospital discharge also increased according to increasing hospitalization severity (**Figure 4** and **Supplemental Table 5**). ICU with ventilation was associated with increased risk of a new mental health condition occurring compared with the outpatient sub-cohort at >30–≤90 days (RR 1.89, 95% CI 1.51, 2.35) and >90–≤180 days post COVID-19 diagnosis or hospital discharge (RR 1.52, 95% CI 1.20, 1.91), and ICU without ventilation was similarly associated with increased risk of a new mental health condition diagnosis during the >90–≤180-day window (RR 1.34, 95% CI 1.02, 1.73). Of note, compared with those <18 years, all age groups examined appeared to be at higher risk of the occurrence of new mental health conditions at >30–≤90 days post COVID-19 diagnosis or hospital discharge. In the >90–

≤180-day window, only the 65–74 and 75–84 years age groups were not at higher risk. Additional covariates associated with increased risk of a new mental health condition occurring included obesity, Caucasian race, and non-Hispanic ethnicity. See **Supplemental Table 5** for full results.

Sensitivity analysis

With the exception of older age, COVID-19 severity did not predict a new cancer diagnosis up to 180 days after COVID-19 diagnosis or hospital discharge (**Supplemental Figure 1** and **Supplemental Table 6**), giving confidence in the results of the original analysis.

Discussion

By utilizing EHRs of over 55,000 patients from hospitals and clinics across the US, this study set out to examine the types of new LTOs (i.e., only those that were identified after COVID-19 diagnosis or hospital discharge) associated with long COVID and to identify potential underlying factors that may contribute to their occurrence. Severe disease was found to predict an increased likelihood of a new LTO diagnosis, whereby increasing hospitalization severity was associated with increased risk of new respiratory (e.g., pneumonia), cardiovascular (e.g., myocardial infarction), and mental health conditions (e.g., confusion or disorientation). In severely affected COVID-19 patients, some LTOs were diagnosed between three and six months after hospital discharge, suggesting that the overall COVID-19 burden extends far beyond the acute infection phase. In addition, although patients with severe disease were most at risk of presenting with new LTOs, non-hospitalized patients also experienced a relatively high incidence of LTOs, suggesting that even patients with mild disease are at risk of adverse long-term effects associated with COVID-19.

Although the data show a clear general trend of increased LTOs that correlated with COVID-19 severity, the specificity of this effect to COVID-19 is unclear, as ICU survivors commonly develop a range of new conditions upon discharge collectively referred to as 'post-intensive care syndrome', regardless of their underlying diagnosis.[11] Nonetheless, preventing the development of more severe disease, where possible, may decrease the likelihood of health problems post infection and would be expected to simultaneously increase the probability of survival. Together, these effects would have a cumulative positive impact on both patients and healthcare systems.

Interestingly, the 'ER on diagnosis' sub-cohort exhibited a reduced incidence of LTOs compared with the outpatient sub-cohort. The reasons for this are not clear but are likely due in part to the lower mean age and reduced incidence of comorbidities in this sub-cohort relative to the other sub-cohorts. In addition, it is possible that in the context of the pandemic, when primary care physicians had more limited personal protective equipment

and other resources, these patients were directed to the ER to be tested for COVID-19, despite not having severe enough disease to warrant an ER visit. Finally, depending on the hospital setting and processes in place, asymptomatic patients who attended the ER for non-COVID-19 reasons may have tested positive while there, which may have led to the inclusion of milder COVID-19 cases in this sub-cohort.

Previous studies have examined the link between COVID-19 severity and LTOs. A study of 2,469 hospitalized COVID-19 patients in Wuhan, China showed that more severe disease correlated with increased risk of LTOs up to six months after infection, including fatigue, sleep difficulties, and anxiety or depression.[23] Anxiety or depression was observed in 23% of patients in that study compared with ~10% in our study; this difference is likely because our study was limited to newly diagnosed disorders in both inpatients and outpatients, while the previous study included new or worsening symptoms in hospitalized patients only. A separate, large study of COVID-19 patients that utilized a US EHR database (N=236,379) to examine six-month outcomes (inpatients and outpatients) reported that ~7% of patients had a first anxiety disorder compared with ~17% that had any anxiety disorder, and that increased incidence was correlated with increased disease severity.[13] A further study compared 73,435 non-hospitalized COVID-19 patients who were users of the Veterans Health Administration with 4,990,835 control patients and reported an increased risk of incident sequelae including, but not limited to, respiratory, cardiovascular, and mental health disorders after a median follow-up duration of 126 and 130 days, respectively.[19] Smaller, single-site hospital studies in the United Kingdom have reported similar trends between disease severity and shorter-term outcomes, with breathlessness commonly reported up to 12 weeks post COVID-19.[24 25] In addition, self-reported data in patients with COVID-19 (N=4,182) showed that upper respiratory complaints (e.g., shortness of breath) and cardiac symptoms (e.g., palpitations, tachycardia) were commonly reported in patients with long COVID (symptoms lasting ≥28 days),[26] and data from a separate study utilizing wearable devices provided further evidence of prolonged tachycardia in symptomatic patients with COVID-19.[27] The current study builds on these previous reports and provides additional

evidence of a link between COVID-19 severity and increased risk of developing LTOs, using a large dataset from both hospitalized and non-hospitalized patients. In addition, our study provides a detailed summary of the incidence of a wide range of specific health conditions that occurred up to six months after COVID-19 diagnosis or hospital discharge, providing a useful resource to better understand and characterise the range of conditions that constitute long COVID.

Our study categorized three major classes of LTOs that occur in patients with long COVID: respiratory, cardiovascular, and mental health. This is broadly in keeping with a previous retrospective cohort study in England that followed 48,780 patients hospitalized with COVID-19, who had significantly higher rates of respiratory and cardiovascular disease after a mean follow-up of 140 days.[28] In addition, a retrospective study that used a large administrative all-payor database including 27,589 inpatients and 46,857 outpatients demonstrated that post COVID-19, patients were more likely to experience a range of conditions, including respiratory, nervous, and circulatory system conditions, than outpatient control patients.[29] A greater understanding of the conditions that characterize long COVID is needed to better anticipate the future healthcare burden of COVID-19 and to optimize strategies to minimize long COVID development. In this regard, signals detected in the current study such as lung fibrosis, as well as other factors including pediatric long COVID, vaccination effects, and healthcare utilization, are topics that may warrant future analysis. In particular, a greater understanding of the long-term economic consequences of COVID-19 and the impact of long COVID on patient quality of life is needed.

A major limitation of this analysis is that treatment setting is used as a proxy for COVID-19 severity; therefore, it is difficult to tease out the effect of treatment setting procedures (e.g., invasive ventilation) from the underlying COVID-19 severity. Furthermore, our analysis did not distinguish short-term from chronic health conditions. Additional limitations include missing information on smoking status, the restriction of follow-up to only six months, the lack of a COVID-19-negative control group, the possibility of missing data

(e.g., patients may have sought care for an LTO not captured in the Optum® de-identified COVID-19 EHR dataset), the lack of information on COVID-19 treatments received, and the lack of laboratory values or other biomarkers to better characterize disease. Finally, capture of health conditions relies on International Classification of Disease-10 (ICD-10) codes, whereas some conditions of interest (e.g., anosmia, ageusia, and brain fog) lack specific ICD-10 codes and other conditions are known to be under-captured.

Conclusions

Although LTOs were reported in patients across all sub-cohorts, increased risk of new respiratory, cardiovascular, and mental health conditions was observed with increasing COVID-19 severity. Strikingly, the risk of new conditions being diagnosed remained high up to six months post COVID-19 diagnosis or hospital discharge, suggesting that the burden of COVID-19 extends far beyond the acute infection phase. Future research is warranted to understand specific factors that lead to the occurrence of new LTOs in patients with COVID-19, and to distinguish between the relative effect of COVID-19 severity versus any general effects that may occur after acute critical illness.

Acknowledgements

We thank Shemra Rizzo for valuable contributions. Third-party medical writing assistance, under the direction of the authors, was provided by John Bett, PhD, of Ashfield MedComms, an Ashfield Health company, and funded by F. Hoffmann-La Roche Ltd.

Contributors

All authors were involved in drafting and revising the manuscript, approved the final version, and agree to being accountable for all aspects of the work. Nick Jovanoski contributed to the

conception of the research question, study design, analysis, and data interpretation. Xin Chen contributed to study design, analysis and data interpretation. Ursula Becker contributed to the conception of the research question, study design, analysis, and data interpretation. Kelly Zalocusky contributed to the conception of the research question, design of the analysis, selection of outcomes and data interpretation. Devika Chawla contributed to the conception of the research question, design of the analysis, and selection of outcomes. Larry Tsai contributed to study design and data interpretation. Michelle Borm contributed to data interpretation. Margaret Neighbors contributed to selection and categorization of key complications for study design. Vincent Yau contributed to the study design, acquisition, analysis and data interpretation.

Funding statement

This work was supported by F. Hoffmann-La Roche Ltd.

Competing interests statement

Nick Jovanoski and Ursula Becker are employees of F. Hoffmann-La Roche Ltd. Michelle Borm is an employee of Roche Nederland BV. Ursula Becker and Michelle Borm hold shares in F. Hoffmann-La Roche Ltd. Xin Chen, Kelly Zalocusky, Devika Chawla, Larry Tsai, Margaret Neighbors, and Vincent Yau are employees of Genentech, Inc. and hold shares in F. Hoffmann-La Roche Ltd.

Patient consent

None required

Ethics approval

The use of the Optum® de-identified COVID-19 EHR dataset was reviewed by the New England Institutional Review Board (IRB) and was determined to be exempt from broad IRB approval, as this study did not involve human subject research.

Data availability statement

Data may be obtained from a third party and are not publicly available. Data were licensed from Optum® and interested researchers may contact Optum® for data access requests. All interested researchers can access the data in the same manner as the authors. The authors had no special access privileges.

References

1. Centers for Disease Control and Prevention. Estimated Disease Burden of COVID-19 2021 [Available from: https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/burden.html accessed May 2021.
2. Carfi A, Bernabei R, Landi F, et al. Persistent Symptoms in Patients After Acute COVID-19. *JAMA* 2020;324(6):603-05. doi: 10.1001/jama.2020.12603 [published Online First: 2020/07/10]
3. Tenforde MW, Kim SS, Lindsell CJ, et al. Symptom Duration and Risk Factors for Delayed Return to Usual Health Among Outpatients with COVID-19 in a Multistate Health Care Systems Network - United States, March-June 2020. *MMWR Morb Mortal Wkly Rep* 2020;69(30):993-98. doi: 10.15585/mmwr.mm6930e1 [published Online First: 2020/07/31]
4. Nalbandian A, Sehgal K, Gupta A, et al. Post-acute COVID-19 syndrome. *Nat Med* 2021;27(4):601-15. doi: 10.1038/s41591-021-01283-z [published Online First: 2021/03/24]
5. Centers for Disease Control and Prevention. Post-COVID Conditions 2021 [Available from: https://www.cdc.gov/coronavirus/2019-ncov/long-term-effects.html accessed May 2021.
6. Augustin M, Schommers P, Stecher M, et al. Post-COVID syndrome in non-hospitalised patients with COVID-19: a longitudinal prospective cohort study. *Lancet Reg Health Eur* 2021;6:100122. doi: 10.1016/j.lanepe.2021.100122 [published Online First: 2021/05/25]
7. Mullard A. Long COVID's long R&D agenda. *Nat Rev Drug Discov* 2021;20(5):329-31. doi: 10.1038/d41573-021-00069-9 [published Online First: 2021/04/22]
8. National Institutes of Health. NIH launches new initiative to study “Long COVID” Bethesda, Maryland: National Institutes of Health; 2021 [Available from: https://www.nih.gov/about-nih/who-we-are/nih-director/statements/nih-launches-new-initiative-study-long-covid accessed July 2021.

9. Mayor N, Tsang R, Joy M, et al. Long covid: coding is caring. *BMJ* 2021;373:n1262. doi: 10.1136/bmj.n1262 [published Online First: 2021/05/21]

10. Marx V. Scientists set out to connect the dots on long COVID. *Nat Methods* 2021;18(5):449-53. doi: 10.1038/s41592-021-01145-z [published Online First: 2021/04/30]

11. Brown SM, Bose S, Banner-Goodspeed V, et al. Approaches to Addressing Post-Intensive Care Syndrome among Intensive Care Unit Survivors. A Narrative Review. *Ann Am Thorac Soc* 2019;16(8):947-56. doi: 10.1513/AnnalsATS.201812-913FR [published Online First: 2019/06/05]

12. Myall KJ, Mukherjee B, Castanheira AM, et al. Persistent Post-COVID-19 Interstitial Lung Disease. An Observational Study of Corticosteroid Treatment. *Ann Am Thorac Soc* 2021;18(5):799-806. doi: 10.1513/AnnalsATS.202008-1002OC [published Online First: 2021/01/13]

13. Taquet M, Geddes JR, Husain M, et al. 6-month neurological and psychiatric outcomes in 236 379 survivors of COVID-19: a retrospective cohort study using electronic health records. *Lancet Psychiatry* 2021;8(5):416-27. doi: 10.1016/S2215-0366(21)00084-5 [published Online First: 2021/04/10]

14. Puntmann VO, Carerj ML, Wieters I, et al. Outcomes of Cardiovascular Magnetic Resonance Imaging in Patients Recently Recovered From Coronavirus Disease 2019 (COVID-19). *JAMA Cardiol* 2020;5(11):1265-73. doi: 10.1001/jamacardio.2020.3557 [published Online First: 2020/07/31]

15. University of Birmingham. Why are some people more susceptible to the long-term health effects of COVID-19? 2021 [Available from: https://www.birmingham.ac.uk/research/quest/21st-century-healthcare/long-covid.aspx accessed May 2021.

16. Mount Sinai Hospital. Center for Post-COVID Care 2021 [Available from: https://www.mountsinai.org/about/covid19/center-post-covid-care accessed May 2021.

17. Chawla D, Rizzo S, Zalocusky K, et al. Descriptive epidemiology of 16,780 hospitalized COVID-19 patients in the United States. *medRxiv* 2021 doi: https://doi.org/10.1101/2020.07.17.20156265

18. Incerti D, Rizzo S, Li X, et al. Prognostic model to identify and quantify risk factors for mortality among hospitalised patients with COVID-19 in the USA. *BMJ open* 2021;11(4):e047121.

19. Al-Aly Z, Xie Y, Bowe B. High-dimensional characterization of post-acute sequelae of COVID-19. *Nature* 2021 doi: 10.1038/s41586-021-03553-9

20. Centers for Disease Control and Prevention. People with Certain Medical Conditions 2021 [Available from: https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/people-with-medical-conditions.html accessed July 2021.

21. R Core Team. R: A language and environment for statistical computing Vienna, Austria2020 [Available from: https://www.r-project.org/.

22. Lüdecke D. Statistical Functions for Regression Models (Version 0.18.1) 2021 [Available from: https://CRAN.R-project.org/package=sjstats.

23. Huang C, Huang L, Wang Y, et al. 6-month consequences of COVID-19 in patients discharged from hospital: a cohort study. *Lancet* 2021;397(10270):220-32. doi: 10.1016/S0140-6736(20)32656-8 [published Online First: 2021/01/12]

24. Halpin SJ, McIvor C, Whyatt G, et al. Postdischarge symptoms and rehabilitation needs in survivors of COVID-19 infection: A cross-sectional evaluation. *J Med Virol* 2021;93(2):1013-22. doi: 10.1002/jmv.26368 [published Online First: 2020/07/31]

25. Arnold DT, Hamilton FW, Milne A, et al. Patient outcomes after hospitalisation with COVID-19 and implications for follow-up: results from a prospective UK cohort. *Thorax* 2020 doi: 10.1136/thoraxjnl-2020-216086 [published Online First: 2020/12/05]

26. Sudre CH, Murray B, Varsavsky T, et al. Attributes and predictors of long COVID. *Nature Medicine* 2021;27(4):626-31. doi: 10.1038/s41591-021-01292-y

27. Radin JM, Quer G, Ramos E, et al. Assessment of Prolonged Physiological and Behavioral Changes Associated With COVID-19 Infection. *JAMA Netw Open*

2021;4(7):e2115959. doi: 10.1001/jamanetworkopen.2021.15959 [published Online First: 2021/07/08]

28. Ayoubkhani D, Khunti K, Nafilyan V, et al. Post-covid syndrome in individuals admitted to hospital with covid-19: retrospective cohort study. *BMJ* 2021;372:n693. doi: 10.1136/bmj.n693 [published Online First: 2021/04/02]
29. Chevinsky JR, Tao G, Lavery AM, et al. Late conditions diagnosed 1-4 months following an initial COVID-19 encounter: a matched cohort study using inpatient and outpatient administrative data - United States, March 1-June 30, 2020. *Clin Infect Dis* 2021 doi: 10.1093/cid/ciab338 [published Online First: 2021/04/29]

Tables and Figures

Table 1 Baseline characteristics of COVID-19 patients overall and by sub-cohort

		Sub-cohort					
	All patients (N=57,748)	**1. Outpatient (n=22,788)**	**2. ER on diagnosis (n=11,633)**	**3. ER (n=2,877)**	**4. Hospitalization without ICU (n=16,653)**	**5. ICU without ventilation (n=1,837)**	**6. ICU with ventilation (n=1,960)**
Mean age (SD), years	47.93 (18.76)	46.78 (18.90)	40.95 (16.89)	44.00 (16.61)	53.17 (18.50)	55.92 (17.32)	57.70 (14.78)
Age group, n (%)							
<18 years	2,184 (3.8)	1,033 (4.5)	641 (5.5)	78 (2.7)	366 (2.2)	46 (2.5)	20 (1.0)
18–29 years	8,509 (14.7)	3,693 (16.2)	2,543 (21.9)	538 (18.7)	1,574 (9.5)	89 (4.8)	72 (3.7)
30–39 years	8,972 (15.5)	3,541 (15.5)	2,496 (21.5)	579 (20.1)	2,046 (12.3)	175 (9.5)	135 (6.9)
40–49 years	9,362 (16.2)	3,664 (16.1)	2,205 (19.0)	555 (19.3)	2,442 (14.7)	253 (13.8)	243 (12.4)
50–64 years	16,103 (27.9)	6,231 (27.3)	2,706 (23.3)	780 (27.1)	4,937 (29.6)	626 (34.1)	823 (42.0)
65–74 years	7,065 (12.2)	2,658 (11.7)	689 (5.9)	230 (8.0)	2,683 (16.1)	363 (19.8)	442 (22.6)
75–84 years	3,620 (6.3)	1,279 (5.6)	239 (2.1)	76 (2.6)	1,647 (9.9)	195 (10.6)	184 (9.4)
≥85 years	891 (1.5)	303 (1.3)	54 (0.5)	22 (0.8)	440 (2.6)	44 (2.4)	28 (1.4)
Missing	1,042 (1.8)	386 (1.7)	60 (0.5)	19 (0.7)	518 (3.1)	46 (2.5)	13 (0.7)
Sex, n (%)							

Female	30,782 (53.3)	12,856 (56.4)	6,115 (52.6)	1,721 (59.8)	8,487 (51.0)	829 (45.1)	774 (39.5)
Male	26,939 (46.6)	9,920 (43.5)	5,515 (47.4)	1,152 (40.0)	8,160 (49.0)	1,008 (54.9)	1,184 (60.4)
Missing	27 (<0.1)	12 (<0.1)	3 (<0.1)	4 (<0.1)	6 (<0.1)	0 (0.0)	2 (<0.1)
Race, n (%)							
African American	13,183 (22.8)	3,473 (15.2)	3,178 (27.3)	790 (27.5)	4,675 (28.1)	551 (30.0)	516 (26.3)
Asian	1,848 (3.2)	639 (2.8)	438 (3.8)	100 (3.5)	555 (3.3)	41 (2.2)	75 (3.8)
Caucasian	29,074 (50.3)	13,746 (60.3)	4,653 (40.0)	1,337 (46.5)	7,538 (45.3)	849 (46.2)	951 (48.5)
Missing	13,643 (23.6)	4,930 (21.6)	3,364 (28.9)	650 (22.6)	3,885 (23.3)	396 (21.6)	418 (21.3)
Ethnicity, n (%)							
Hispanic	11,932 (20.7)	3,942 (17.3)	3,378 (29.0)	646 (22.5)	3,298 (19.8)	332 (18.1)	336 (17.1)
Non-Hispanic	38,988 (67.5)	15,485 (68.0)	7,121 (61.2)	1,987 (69.1)	11,648 (69.9)	1,294 (70.4)	1,453 (74.1)
Missing	6,828 (11.8)	3,361 (14.7)	1,134 (9.7)	244 (8.5)	1,707 (10.3)	211 (11.5)	171 (8.7)
Smoking status, n (%)							
Current smoker	413 (0.7)	193 (0.8)	128 (1.1)	13 (0.5)	61 (0.4)	13 (0.7)	5 (0.3)
Previously smoked	740 (1.3)	417 (1.8)	111 (1.0)	27 (0.9)	145 (0.9)	25 (1.4)	15 (0.8)
Never smoked	2,831 (4.9)	1,468 (6.4)	750 (6.4)	121 (4.2)	404 (2.4)	50 (2.7)	38 (1.9)
Missing	53,764 (93.1)	20,710 (90.9)	10,644 (91.5)	2,716 (94.4)	16,043 (96.3)	1,749 (95.2)	1,902 (97.0)

Obese, n (%)*	10,952 (19.0)	4,905 (21.5)	1,366 (11.7)	580 (20.2)	3,246 (19.5)	406 (22.1)	449 (22.9)
Insurance, n (%)							
Commercial	29,145 (50.5)	13,134 (57.6)	5,672 (48.8)	1,482 (51.5)	7,243 (43.5)	758 (41.3)	856 (43.7)
Medicaid	8,652 (15.0)	2,341 (10.3)	2,223 (19.1)	542 (18.8)	2,891 (17.4)	312 (17.0)	343 (17.5)
Medicare	8,774 (15.2)	3,173 (13.9)	788 (6.8)	245 (8.5)	3,674 (22.1)	435 (23.7)	459 (23.4)
Other payor type	4,004 (6.9)	1,282 (5.6)	1,071 (9.2)	211 (7.3)	1,188 (7.1)	129 (7.0)	123 (6.3)
Uninsured	4,833 (8.4)	1,542 (6.8)	1,731 (14.9)	281 (9.8)	1,069 (6.4)	111 (6.0)	99 (5.1)
Missing	2,340 (4.1)	1,316 (5.8)	148 (1.3)	116 (4.0)	588 (3.5)	92 (5.0)	80 (4.1)
Month of COVID-19 diagnosis, n (%)							
Feb 2020	115 (0.2)	61 (0.3)	3 (<0.1)	4 (0.1)	34 (0.2)	5 (0.3)	8 (0.4)
Mar 2020	8,197 (14.2)	1,527 (6.7)	1,893 (16.3)	527 (18.3)	3,288 (19.7)	306 (16.7)	656 (33.5)
Apr 2020	18,591 (32.2)	6,018 (26.4)	3,480 (29.9)	926 (32.2)	6,684 (40.1)	676 (36.8)	807 (41.2)
May 2020	14,188 (24.6)	6,154 (27.0)	2,703 (23.2)	729 (25.3)	3,755 (22.5)	477 (26.0)	370 (18.9)
Jun 2020	14,832 (25.7)	7,846 (34.4)	3,073 (26.4)	597 (20.8)	2,826 (17.0)	371 (20.2)	119 (6.1)
Jul 2020	1,825 (3.2)	1,182 (5.2)	481 (4.1)	94 (3.3)	66 (0.4)	2 (0.1)	0 (0.0)
Mean weighted CCI (SD)	1.20 (2.06)	1.09 (2.03)	0.56 (1.32)	0.88 (1.76)	1.64 (2.31)	2.15 (2.46)	2.13 (2.39)

CCI, Charlson Comorbidity Index; COVID-19, coronavirus disease 2019; ER, emergency room; ICU, intensive care unit; SD, standard deviation

*No patient data were missing for obesity

Table 2 Long-term outcomes that were diagnosed >30–≤90 days or >90–≤180 days post COVID-19 by sub-cohort

Condition, n **(%)**	**1. Outpatient (N=22,788)**		**2. ER on diagnosis (N=11,633)**		**3. ER (N=2,877)**		**4. Hospitalization without ICU (N=16,653)**		**5. ICU without ventilation (N=1,837)**		**6. ICU with ventilation (N=1,960)**	
	>30–≤90 days	**>90–≤180 days**	**>30–≤90 days**	**>90–≤180 days**	**>30–≤90 days**	**>90–≤180 days**	**>30–≤90 days**	**>90–≤180 days**	**>30–≤90 days**	**>90–≤180 days**	**>30–≤90 days**	**>90–≤180 days**
Respiratory	1,846 (**8.1**)	1,846 (**8.1**)	520 (**4.5**)	484 (**4.2**)	322 (**11.2**)	311 (**10.8**)	2,124 (**12.8**)	1,599 (**9.6**)	296 (**16.1**)	222 (**12.1**)	547 (**27.9**)	381 (**19.4**)
CV	439 (**1.9**)	475 (**2.1**)	81 (**0.7**)	92 (**0.8**)	54 (**1.9**)	57 (**2.0**)	585 (**3.5**)	618 (**3.7**)	102 (**5.6**)	99 (**5.4**)	128 (**6.5**)	132 (**6.7**)
Mental health	940 (**4.1**)	1,081 (**4.7**)	231 (**2.0**)	293 (**2.5**)	144 (**5.0**)	154 (**5.4**)	819 (**4.9**)	866 (**5.2**)	119 (**6.5**)	118 (**6.4**)	167 (**8.5**)	144 (**7.3**)
Cancer	137 (**0.6**)	151 (**0.7**)	24 (**0.2**)	30 (**0.3**)	18 (**0.6**)	14 (**0.5**)	95 (**0.6**)	106 (**0.6**)	9 (**0.5**)	12 (**0.7**)	22 (**1.1**)	14 (**0.7**)
Respiratory and CV	131 (**0.6**)	124 (**0.5**)	16 (**0.1**)	30 (**0.3**)	17 (**0.6**)	19 (**0.7**)	201 (**1.2**)	181 (**1.1**)	30 (**1.6**)	26 (**1.4**)	63 (**3.2**)	47 (**2.4**)
Respiratory and mental health	205 (**0.9**)	216 (**0.9**)	65 (**0.6**)	63 (**0.5**)	48 (**1.7**)	33 (**1.1**)	262 (**1.6**)	245 (**1.5**)	45 (**2.4**)	34 (**1.9**)	83 (**4.2**)	50 (**2.6**)
Mental health and CV	53 (**0.2**)	49 (**0.2**)	8 (**0.1**)	9 (**0.1**)	4 (**0.1**)	2 (**0.1**)	57 (**0.3**)	64 (**0.4**)	10 (**0.5**)	11 (**0.6**)	7 (**0.4**)	11 (**0.6**)
Respiratory, CV, and mental health	57 (**0.3**)	52 (**0.2**)	9 (**0.1**)	10 (**0.1**)	6 (**0.2**)	9 (**0.3**)	98 (**0.6**)	84 (**0.5**)	17 (**0.9**)	15 (**0.8**)	30 (**1.5**)	31 (**1.6**)
No new conditions* (respiratory, CV, or mental health)	20,066 (**88.1**)	19,879 (**87.2**)	10,908 (**93.8**)	10,886 (**93.6**)	2,438 (**84.7**)	2,427 (**84.4**)	13,841 (**83.1**)	14,228 (**85.4**)	1,439 (**78.3**)	1,499 (**81.6**)	1,331 (**67.9**)	1,473 (**75.2**)

COVID-19, coronavirus disease 2019; CV, cardiovascular; ER, emergency room; ICU, intensive care unit

*Only conditions that appeared >30–≤180 days after COVID-19 diagnosis or hospital discharge are included; pre-existing conditions are excluded

Figure 1

Title: Overall study design.

Abbreviations: COVID-19, coronavirus disease 2019

Figure 2

Title: Relative risk of new respiratory conditions occurring from >30–≤180 days after COVID-19 diagnosis or hospital discharge.

Abbreviations: CCI, Charlson Comorbidity Index; COVID-19, coronavirus disease 2019; ER, emergency room; ICU, intensive care unit

Legend: Relative risk of new respiratory conditions occurring at (A) >30–≤90 days and (B) >90–≤180 days after COVID-19 diagnosis or hospital discharge. Graphs represent relative risk and 95% confidence intervals. Reference groups are: <18 years (age), female (sex), African American (race), Hispanic (ethnicity), non-obese (obesity), diagnosis in February and March 2020 (diagnosis month); commercial (insurance), cohort 1: outpatient (cohort).

Figure 3

Title: Relative risk of new cardiovascular conditions occurring from >30–≤180 days after COVID-19 diagnosis or hospital discharge.

Abbreviations: CCI, Charlson Comorbidity Index; COVID-19, coronavirus disease 2019; ER, emergency room; ICU, intensive care unit

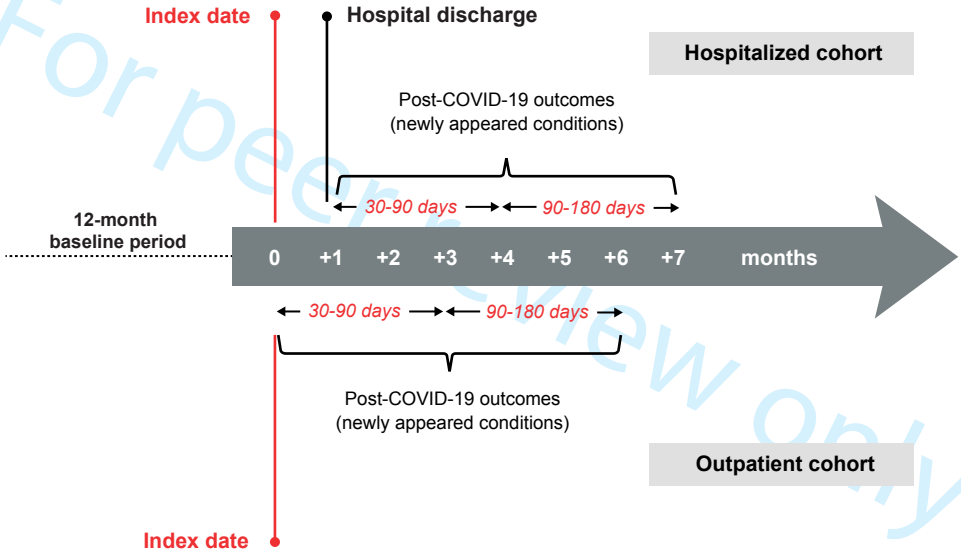
Legend: Relative risk of new cardiovascular conditions occurring at (A) >30–≤90 days and (B) >90–≤180 days after COVID-19 diagnosis or hospital discharge. Graphs represent relative risk and 95% confidence intervals. Reference groups are: <18 years (age), female (sex), African American (race), Hispanic (ethnicity), non-obese (obesity), diagnosis in February and March 2020 (diagnosis month); commercial (insurance), cohort 1: outpatient (cohort). Relative risk in the >30–≤90 days time window was not calculated as no new diagnoses were made in the reference group (<18 years) during this time.

Figure 4

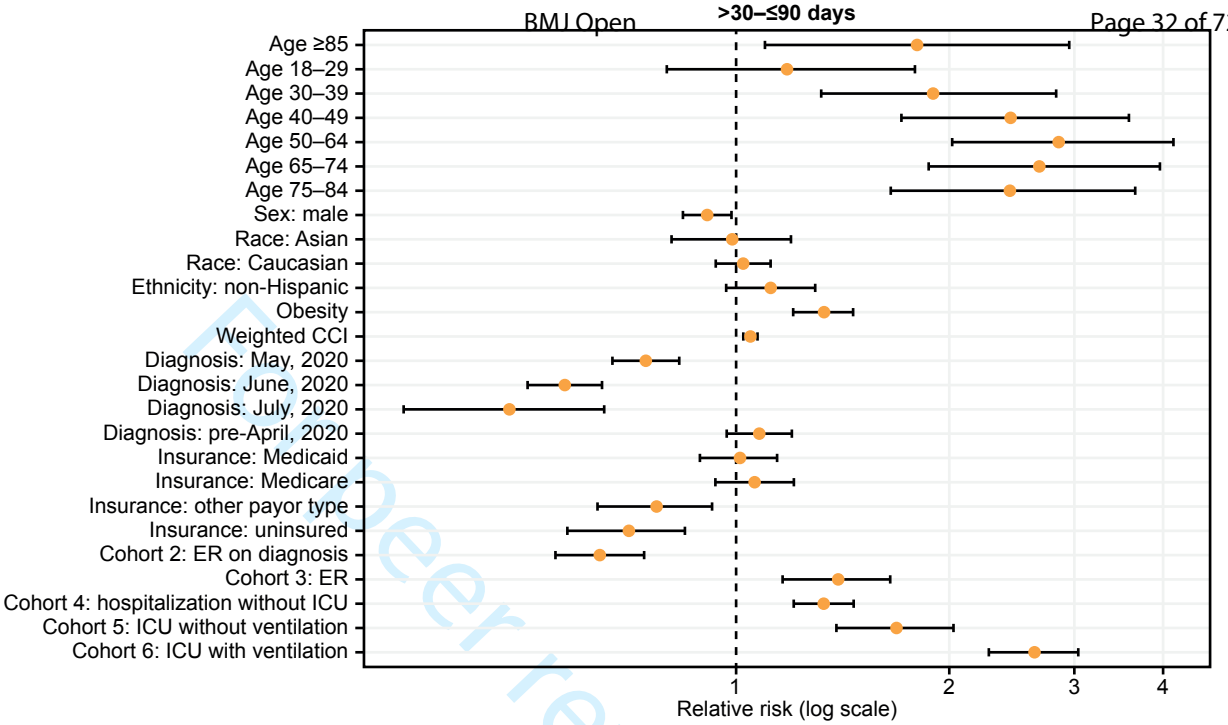
Title: Relative risk of new mental health conditions occurring from >30–≤180 days after COVID-19 diagnosis or hospital discharge.

Abbreviations: CCI, Charlson Comorbidity Index; COVID-19, coronavirus disease 2019; ER, emergency room; ICU, intensive care unit

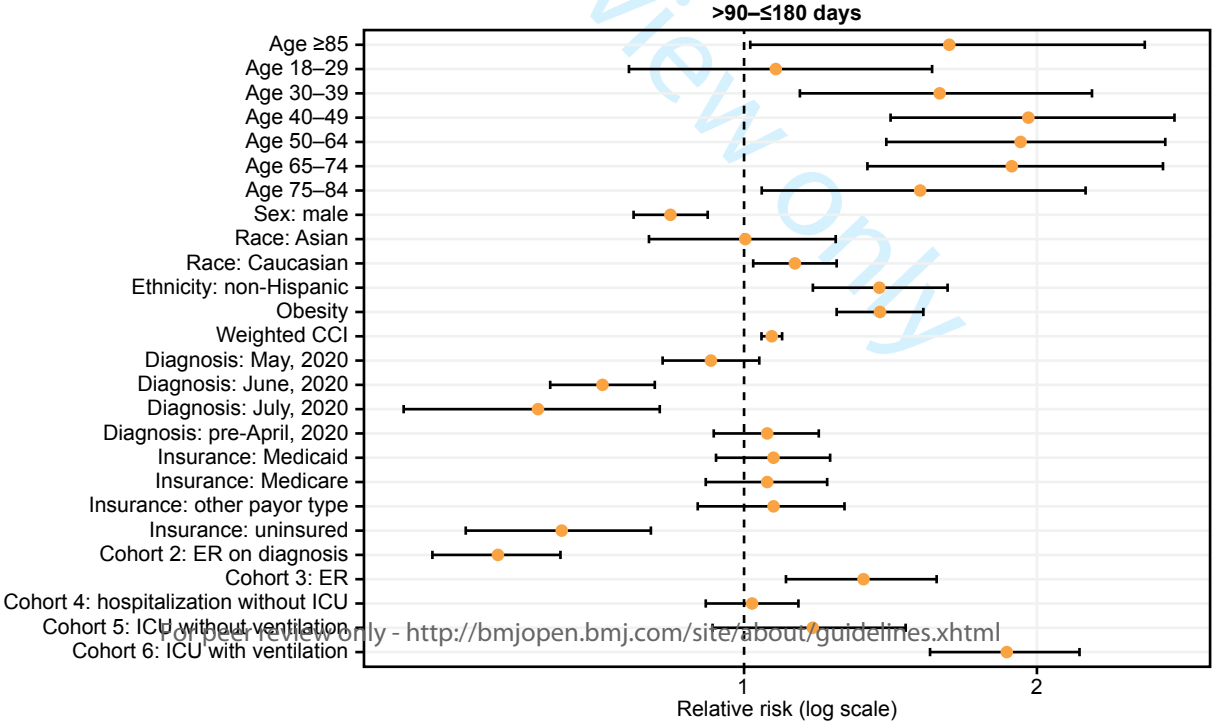
Legend: Relative risk of new mental health conditions occurring at (A) >30–≤90 days and (B) >90–≤180 days after COVID-19 diagnosis or hospital discharge. Graphs represent relative risk and 95% confidence intervals. Reference groups are: <18 years (age), female (sex), African American (race), Hispanic (ethnicity), non-obese (obesity), diagnosis in February and March 2020 (diagnosis month); commercial (insurance), cohort 1: outpatient (cohort).

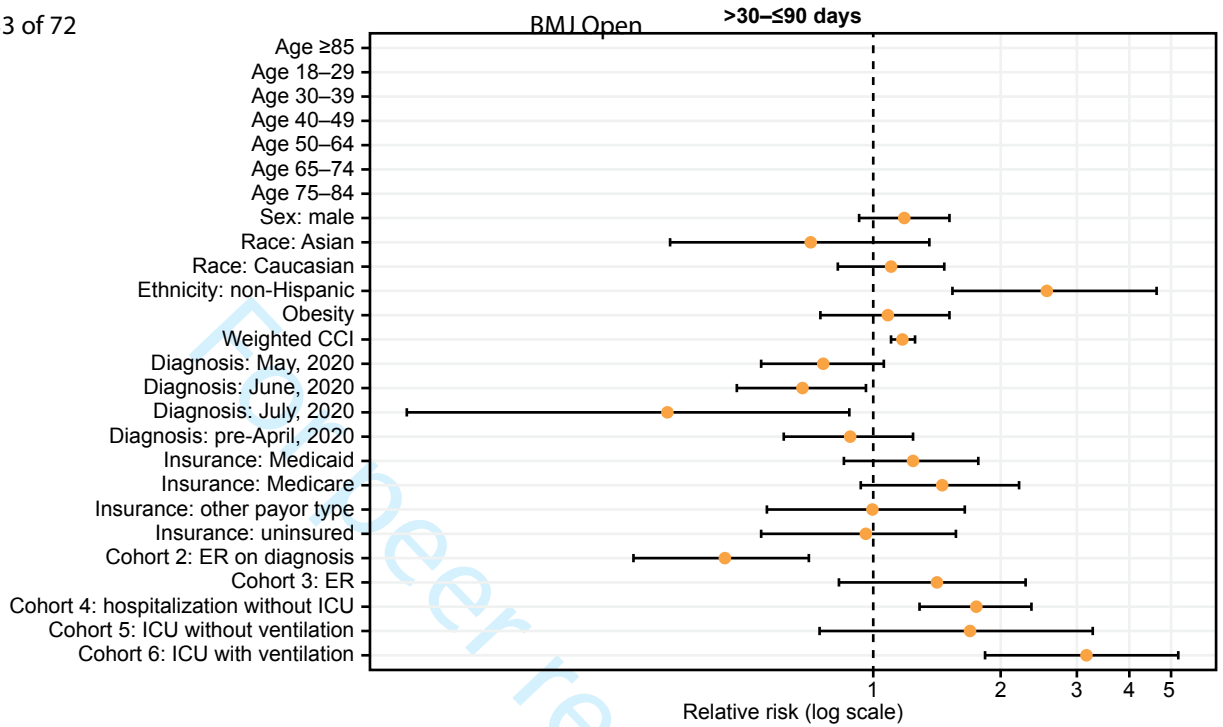


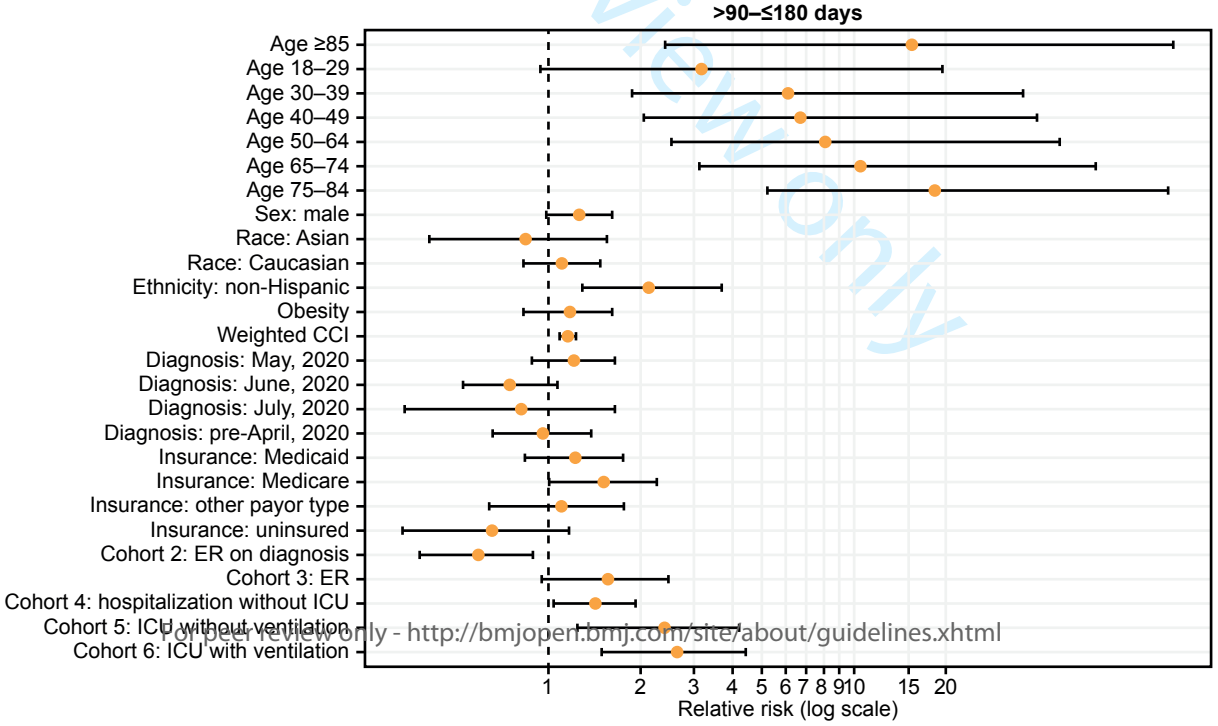
A

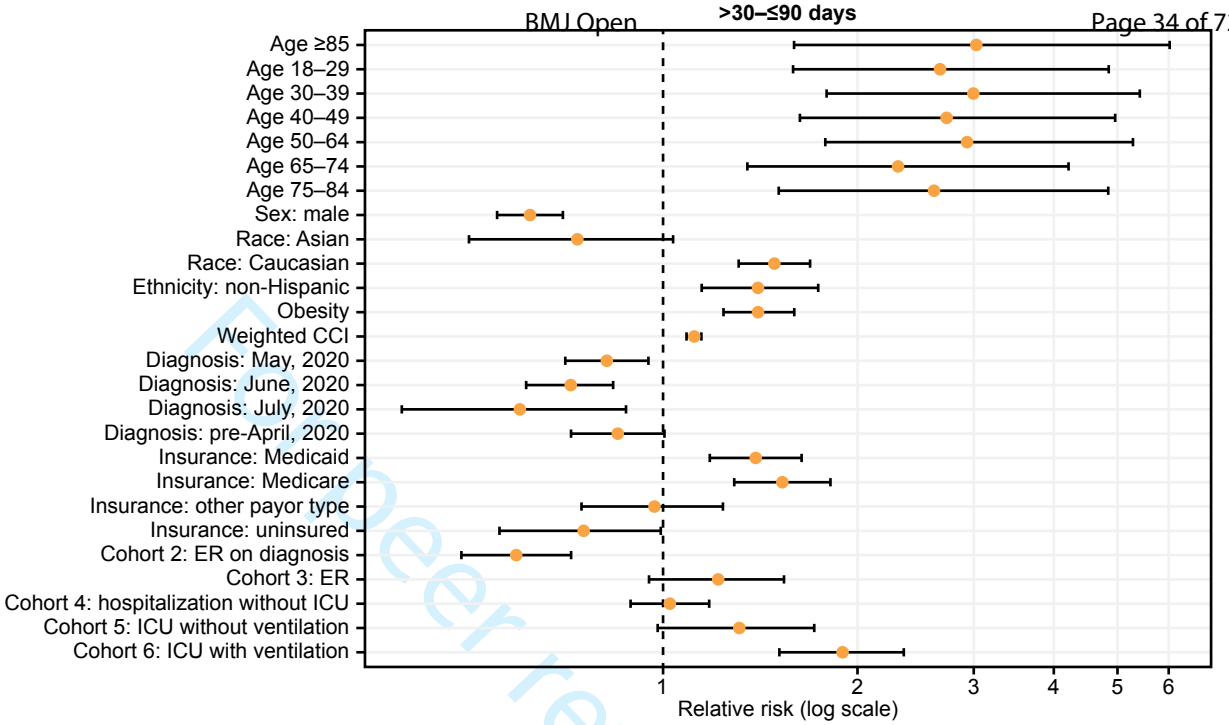


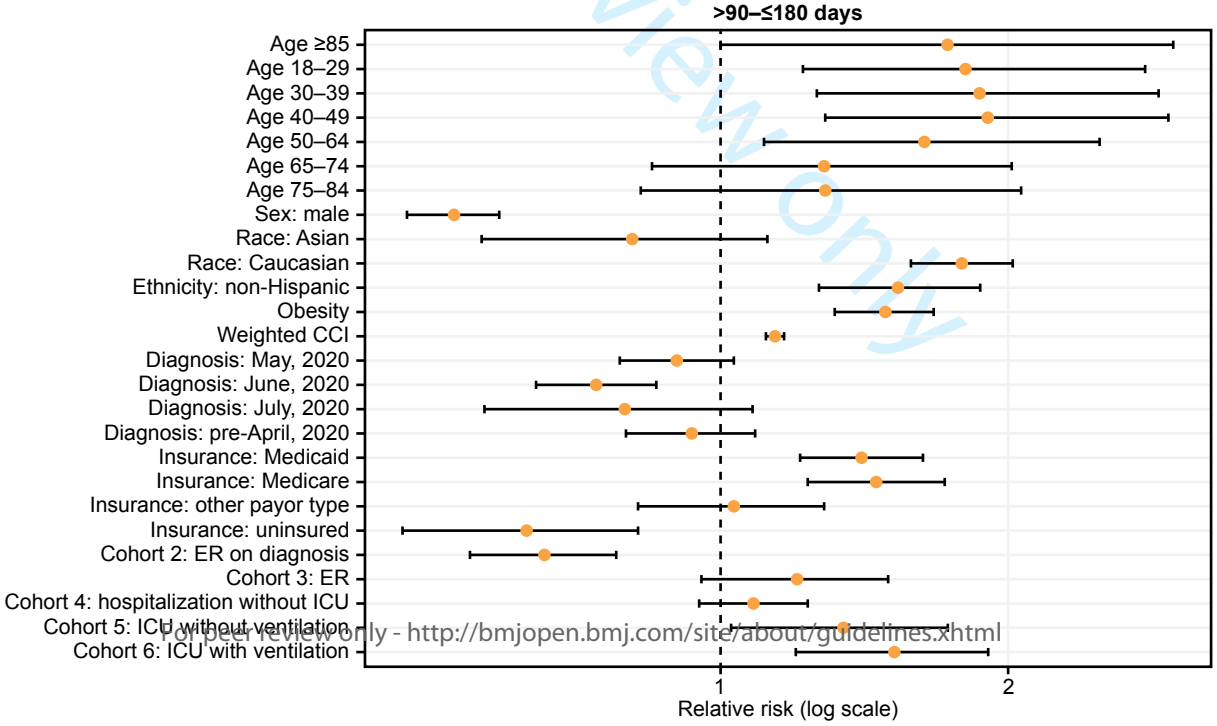
B







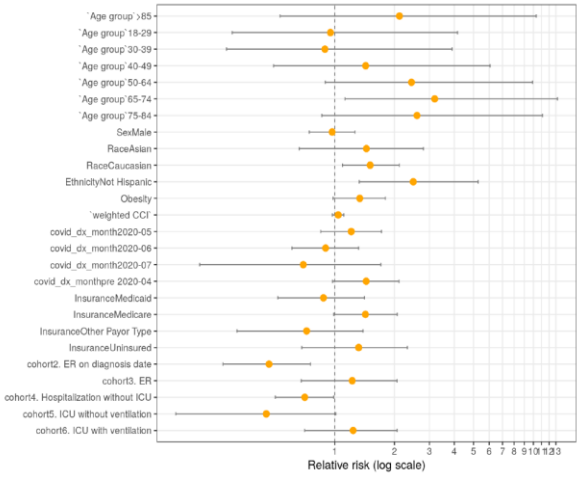




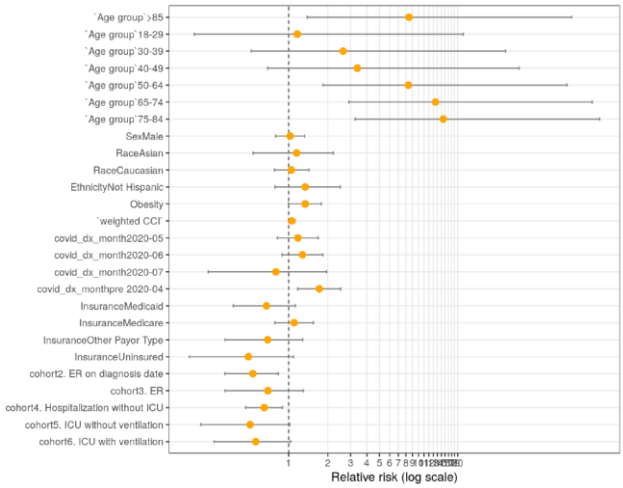
Supplemental materials

Supplemental Figure 1 Relative risk of a new cancer diagnosis from >30–≤180 days after COVID-19 diagnosis or hospital discharge.

A



B



CCI, Charlson Comorbidity Index; COVID-19, coronavirus disease 2019; ER, emergency room; ICU, intensive care unit

Relative risk of a new cancer diagnosis at (A) >30–≤90 days and (B) >90–≤180 days after COVID-19 diagnosis or hospital discharge. Graphs represent relative risk and 95% confidence intervals. Reference groups are: <18 years (age), female (sex), African American (race), Hispanic (ethnicity), non-obese (obesity), diagnosis in February and March 2020 (diagnosis month); commercial (insurance), cohort 1: outpatient (cohort).

Supplemental Table 1 List of the long-term outcomes studied and their classification

Long-term outcome
Respiratory
Asthma
Bronchiectasis
Bronchitis
COPD
Dyspnea
Emphysema
Influenza
Interstitial lung disease (fibrosis)
Pneumonia
Respiratory failure
Cardiovascular
Cardiac arrhythmia
Myocardial infarction
Pulmonary embolism
Pulmonary hypertension
Stroke
Mental health
Anxiety
Confusion or disorientation
Dementia
Depression
Encephalopathy
Memory loss

COPD, chronic obstructive pulmonary disease

Supplemental Table 2 List of comorbidities included in the Charlson Comorbidity Index

Comorbidity	
AIDS	Metastatic solid tumor
Cancer	Mild liver disease
Cerebrovascular disease	Moderate or severe liver disease
Chronic pulmonary disease	Moderate or severe renal disease
Congestive heart failure	Myocardial infarction
Dementia	Peptic ulcer disease
Diabetes with complication	Peripheral vascular disease
Diabetes without complication	Rheumatics
Hemiplegia	

AIDS, acquired immunodeficiency syndrome

Supplemental Table 3 Full baseline characteristics, COVID-19-related outcomes, symptoms, and tests

	All patients (N=57,748)	1. Outpatient (n=22,788)	2. ER on diagnosis (n=11,633)	3. ER (n=2,877)	4. Hospitalization without ICU (N=16,653)	5. ICU without ventilation (N=1,837)	6. ICU with ventilation (N=1,960)
Mean age (SD), years	47.93 (18.76)	46.78 (18.90)	40.95 (16.89)	44.00 (16.61)	53.17 (18.50)	55.92 (17.32)	57.70 (14.78)
Age group, n (%)							
<18 years	2,184 (3.8)	1,033 (4.5)	641 (5.5)	78 (2.7)	366 (2.2)	46 (2.5)	20 (1.0)
18–29 years	8,509 (14.7)	3,693 (16.2)	2,543 (21.9)	538 (18.7)	1574 (9.5)	89 (4.8)	72 (3.7)
30–39 years	8,972 (15.5)	3,541 (15.5)	2,496 (21.5)	579 (20.1)	2,046 (12.3)	175 (9.5)	135 (6.9)
40–49 years	9,362 (16.2)	3,664 (16.1)	2,205 (19.0)	555 (19.3)	2,442 (14.7)	253 (13.8)	243 (12.4)
50–64 years	16,103 (27.9)	6,231 (27.3)	2,706 (23.3)	780 (27.1)	4,937 (29.6)	626 (34.1)	823 (42.0)
65–74 years	7,065 (12.2)	2,658 (11.7)	689 (5.9)	230 (8.0)	2,683 (16.1)	363 (19.8)	442 (22.6)
75–84 years	3,620 (6.3)	1,279 (5.6)	239 (2.1)	76 (2.6)	1,647 (9.9)	195 (10.6)	184 (9.4)
≥85 years	891 (1.5)	303 (1.3)	54 (0.5)	22 (0.8)	440 (2.6)	44 (2.4)	28 (1.4)

Missing	1,042 (1.8)	386 (1.7)	60 (0.5)	19 (0.7)	518 (3.1)	46 (2.5)	13 (0.7)
Sex, n (%)							
Female	30,782 (53.3)	12,856 (56.4)	6,115 (52.6)	1,721 (59.8)	8,487 (51.0)	829 (45.1)	774 (39.5)
Male	26,939 (46.6)	9,920 (43.5)	5,515 (47.4)	1,152 (40.0)	8,160 (49.0)	1,008 (54.9)	1,184 (60.4)
Missing	27 (<0.1)	12 (<0.1)	3 (<0.1)	4 (<0.1)	6 (<0.1)	0 (0.0)	2 (<0.1)
Race, n (%)							
African American	13,183 (22.8)	3,473 (15.2)	3,178 (27.3)	790 (27.5)	4,675 (28.1)	551 (30.0)	516 (26.3)
Asian	1,848 (3.2)	639 (2.8)	438 (3.8)	100 (3.5)	555 (3.3)	41 (2.2)	75 (3.8)
Caucasian	29,074 (50.3)	13,746 (60.3)	4,653 (40.0)	1,337 (46.5)	7,538 (45.3)	849 (46.2)	951 (48.5)
Missing	13,643 (23.6)	4,930 (21.6)	3,364 (28.9)	650 (22.6)	3,885 (23.3)	396 (21.6)	418 (21.3)
Ethnicity, n (%)							
Hispanic	11,932 (20.7)	3,942 (17.3)	3,378 (29.0)	646 (22.5)	3,298 (19.8)	332 (18.1)	336 (17.1)
Non-Hispanic	38,988 (67.5)	15,485 (68.0)	7,121 (61.2)	1,987 (69.1)	11,648 (69.9)	1,294 (70.4)	1,453 (74.1)
Missing	6,828 (11.8)	3,361 (14.7)	1,134 (9.7)	244 (8.5)	1,707 (10.3)	211 (11.5)	171 (8.7)

Region, n (%)							
Midwest	22,133 (38.3)	8,137 (35.7)	5,250 (45.1)	1,364 (47.4)	5,686 (34.1)	830 (45.2)	866 (44.2)
Northwest	20,671 (35.8)	8,018 (35.2)	3,261 (28.0)	875 (30.4)	7,375 (44.3)	496 (27.0)	646 (33.0)
South	8,548 (14.8)	3,463 (15.2)	2,004 (17.2)	367 (12.8)	2,212 (13.3)	271 (14.8)	231 (11.8)
West	4,430 (7.7)	2,379 (10.4)	673 (5.8)	169 (5.9)	873 (5.2)	178 (9.7)	158 (8.1)
Missing	1,966 (3.4)	791 (3.5)	445 (3.8)	102 (3.5)	507 (3.0)	62 (3.4)	59 (3.0)
Division, n (%)							
East North Central	15,381 (26.6)	4,833 (21.2)	3,822 (32.9)	982 (34.1)	4,536 (27.2)	551 (30.0)	657 (33.5)
East South Central	1,769 (3.1)	664 (2.9)	404 (3.5)	62 (2.2)	472 (2.8)	124 (6.8)	43 (2.2)
Middle Atlantic	15,163 (26.3)	6,516 (28.6)	1,718 (14.8)	527 (18.3)	5,622 (33.8)	338 (18.4)	442 (22.6)
Mountain	2,221 (3.8)	1,281 (5.6)	257 (2.2)	48 (1.7)	457 (2.7)	78 (4.2)	100 (5.1)
New England	5,478 (9.5)	1,491 (6.5)	1,538 (13.2)	345 (12.0)	1,744 (10.5)	158 (8.6)	202 (10.3)
Other/unknown	2,085 (3.6)	847 (3.7)	467 (4.0)	112 (3.9)	532 (3.2)	64 (3.5)	63 (3.2)
Pacific	2,205 (3.8)	1,096 (4.8)	416 (3.6)	121 (4.2)	414 (2.5)	100 (5.4)	58 (3.0)

South Atlantic/ West South Central	6,767 (11.7)	2,792 (12.3)	1,599 (13.7)	303 (10.5)	1,738 (10.4)	147 (8.0)	188 (9.6)
West North Central	6,679 (11.6)	3,268 (14.3)	1,412 (12.1)	377 (13.1)	1,138 (6.8)	277 (15.1)	207 (10.6)
Smoking status, n (%)							
Current smoker	413 (0.7)	193 (0.8)	128 (1.1)	13 (0.5)	61 (0.4)	13 (0.7)	5 (0.3)
Previously smoked	740 (1.3)	417 (1.8)	111 (1.0)	27 (0.9)	145 (0.9)	25 (1.4)	15 (0.8)
Never smoked	2,831 (4.9)	1,468 (6.4)	750 (6.4)	121 (4.2)	404 (2.4)	50 (2.7)	38 (1.9)
Missing	53,764 (93.1)	20,710 (90.9)	10,644 (91.5)	2,716 (94.4)	16,043 (96.3)	1,749 (95.2)	1,902 (97.0)
Obese, n (%)							
No	47,796 (81.0)	17,883 (78.5)	10,267 (88.3)	2,297 (79.8)	13,407 (80.5)	1,431 (77.9)	1,511 (77.1)
Yes	10,952 (19.0)	4,905 (21.5)	1,366 (11.7)	580 (20.2)	3,246 (19.5)	406 (22.1)	449 (22.9)
Pregnant n (%)							
No	56,137 (97.2)	22,246 (97.6)	11,409 (98.1)	2,802 (97.4)	15,931 (95.7)	1,813 (98.7)	1,936 (98.8)
Yes	1,611 (2.8)	542 (2.4)	224 (1.9)	75 (2.6)	722 (4.3)	24 (1.3)	24.2 (1.2)
Insurance, n (%)							

Commercial	29,145 (50.5)	13,134 (57.6)	5,672 (48.8)	1,482 (51.5)	7,243 (43.5)	758 (41.3)	856 (43.7)
Medicaid	8,652 (15.0)	2,341 (10.3)	2,223 (19.1)	542 (18.8)	2,891 (17.4)	312 (17.0)	343 (17.5)
Medicare	8,774 (15.2)	3,173 (13.9)	788 (6.8)	245 (8.5)	3674 (22.1)	435 (23.7)	459 (23.4)
Other payor type	4,004 (6.9)	1,282 (5.6)	1,071 (9.2)	211 (7.3)	1,188 (7.1)	129 (7.0)	123 (6.3)
Uninsured	4,833 (8.4)	1,542 (6.8)	1,731 (14.9)	281 (9.8)	1,069 (6.4)	111 (6.0)	99 (5.1)
Missing	2,340 (4.1)	1,316 (5.8)	148 (1.3)	116 (4.0)	588 (3.5)	92 (5.0)	80 (4.1)
Month of COVID-19 diagnosis, n (%)							
Feb 2020	115 (0.2)	61 (0.3)	3 (<0.1)	4 (0.1)	34 (0.2)	5 (0.3)	8 (0.4)
Mar 2020	8,197 (14.2)	1,527 (6.7)	1,893 (16.3)	527 (18.3)	3,288 (19.7)	306 (16.7)	656 (33.5)
Apr 2020	18,591 (32.2)	6,018 (26.4)	3,480 (29.9)	926 (32.2)	6,684 (40.1)	676 (36.8)	807 (41.2)
May 2020	14,188 (24.6)	6,154 (27.0)	2,703 (23.2)	729 (25.3)	3755 (22.5)	477 (26.0)	370 (18.9)
Jun 2020	14,832 (25.7)	7,846 (34.4)	3,073 (26.4)	597 (20.8)	2,826 (17.0)	371 (20.2)	119 (6.1)
Jul 2020	1,825 (3.2)	1,182 (5.2)	481 (4.1)	94 (3.3)	66 (0.4)	2 (0.1)	0 (0.0)
Diabetes with complication, n (%)							

No	53,804 (93.2)	21,658 (95.0)	11,356 (97.6)	2,774 (96.4)	14,812 (88.9)	1,553 (84.5)	1,651 (84.2)
Yes	3,944 (6.8)	1,130 (5.0)	277 (2.4)	103 (3.6)	1,841 (11.1)	284 (15.5)	309 (15.8)
Congestive heart failure, n (%)							
No	54,048 (93.6)	21,702 (95.2)	11,430 (98.3)	2,800 (97.3)	14,938 (89.7)	1,553 (84.5)	1,625 (82.9)
Yes	3,700 (6.4)	1,086 (4.8)	203 (1.7)	77 (2.7)	1,715 (10.3)	284 (15.5)	335 (17.1)
Cerebrovascular disease, n (%)							
No	55,258 (95.7)	21,957 (96.4)	11,485 (98.7)	2,811 (97.7)	15,547 (93.4)	1,662 (90.5)	1,796 (91.6)
Yes	2,490 (4.3)	831 (3.6)	148 (1.3)	66 (2.3)	1,106 (6.6)	175 (9.5)	164 (8.4)
Moderate or severe renal disease, n (%)							
No	53,066 (91.9)	21,454 (94.1)	11,387 (97.9)	2,759 (95.9)	14,387 (86.4)	1,497 (81.5)	1,582 (80.7)
Yes	4,682 (8.1)	1,334 (5.9)	246 (2.1)	118 (4.1)	2,266 (13.6)	340 (18.5)	378 (19.3)
Diabetes without complication, n (%)							
No	47,489 (82.2)	19,807 (86.9)	10,435 (89.7)	2,522 (87.7)	12,313 (73.9)	1,184 (64.5)	1,228 (62.7)

Yes	10,259 (17.8)	2,981 (13.1)	1,198 (10.3)	355 (12.3)	4,340 (26.1)	653 (35.5)	732 (37.3)
Chronic pulmonary disease, n (%)							
No	47,794 (82.8)	19,225 (84.4)	10,203 (87.7)	2,359 (82.0)	13,145 (78.9)	1,439 (78.3)	1,423 (72.6)
Yes	9,954 (17.2)	3,563 (15.6)	1,430 (12.3)	518 (18.0)	3,508 (21.1)	398 (21.7)	537 (27.4)
Mild liver disease, n (%)							
No	55,817 (96.7)	22,095 (97.0)	11,441 (98.3)	2,788 (96.9)	15,890 (95.4)	1,746 (95.0)	1,857 (94.7)
Yes	1,931 (3.3)	693 (3.0)	192 (1.7)	89 (3.1)	763 (4.6)	91 (5.0)	103 (5.3)
Peripheral vascular disease, n (%)							
No	55,049 (95.3)	21,773 (95.5)	11,461 (98.5)	2,808 (97.6)	15,516 (93.2)	1,674 (91.1)	1,817 (92.7)
Yes	2,699 (4.7)	1,015 (4.5)	172 (1.5)	69 (2.4)	1,137 (6.8)	163 (8.9)	143 (7.3)
Cancer, n (%)							
No	53,687 (93.0)	20,822 (91.4)	11,198 (96.3)	2,686 (93.4)	15,465 (92.9)	1,689 (91.9)	1,827 (93.2)
Yes	4,061 (7.0)	1,966 (8.6)	435 (3.7)	191 (6.6)	1,188 (7.1)	148 (8.1)	133 (6.8)

Myocardial infarction, n (%)							
No	55,528 (96.2)	22,273 (97.7)	11,459 (98.5)	2,828 (98.3)	15,528 (93.2)	1,659 (90.3)	1,781 (90.9)
Yes	2,220 (3.8)	515 (2.3)	174 (1.5)	49 (1.7)	1,125 (6.8)	178 (9.7)	179 (9.1)
Dementia, n (%)							
No	55,833 (96.7)	22,213 (97.5)	11,541 (99.2)	2,843 (98.8)	15,663 (94.1)	1,697 (92.4)	1,876 (95.7)
Yes	1,915 (3.3)	575 (2.5)	92 (0.8)	34 (1.2)	990 (5.9)	140 (7.6)	84 (4.3)
Peptic ulcer disease, n (%)							
No	57,305 (99.2)	22,620 (99.3)	11,599 (99.7)	2,864 (99.5)	16,494 (99.0)	1,812 (98.6)	1,916 (97.8)
Yes	443 (0.8)	168 (0.7)	34 (0.3)	13 (0.5)	159 (1.0)	25 (1.4)	44 (2.2)
Hemiplegia, n (%)							
No	57,192 (99.0)	22,647 (99.4)	11,596 (99.7)	2,870 (99.8)	16,402 (98.5)	1,783 (97.1)	1,894 (96.6)
Yes	556 (1.0)	141 (0.6)	37 (0.3)	7 (0.2)	251 (1.5)	54 (2.9)	66 (3.4)

Rheumatics, n (%)							
No	56,674 (98.1)	22,348 (98.1)	11,515 (99.0)	2,831 (98.4)	16,285 (97.8)	1,790 (97.4)	1,905 (97.2)
Yes	1,074 (1.9)	440 (1.9)	118 (1.0)	46 (1.6)	368 (2.2)	47 (2.6)	55 (2.8)
Metastatic solid tumor, n (%)							
No	57,146 (99.0)	22,460 (98.6)	11,601 (99.7)	2,858 (99.3)	16,472 (98.9)	1,810 (98.5)	1,945 (99.2)
Yes	602 (1.0)	328 (1.4)	32 (0.3)	19 (0.7)	181 (1.1)	27 (1.5)	15 (0.8)
Moderate or severe liver disease, n (%)							
No	57,495 (99.6)	22,703 (99.6)	11,625 (99.9)	2,872 (99.8)	16,547 (99.4)	1,818 (99.0)	1,930 (98.5)
Yes	253 (0.4)	85 (0.4)	8 (0.1)	5 (0.2)	106 (0.6)	19 (1.0)	30 (1.5)
AIDS, n (%)							
No	56,640 (98.1)	22,229 (97.5)	11,500 (98.9)	2,813 (97.8)	1,6376 (98.3)	1,794 (97.7)	1,928 (98.4)
Yes	1,108 (1.9)	559 (2.5)	133 (1.1)	64 (2.2)	277 (1.7)	43 (2.3)	32 (1.6)
Mean weighted CCI (SD)	1.20 (2.06)	1.09 (2.03)	0.56 (1.32)	0.88 (1.76)	1.64 (2.31)	2.15 (2.46)	2.13 (2.39)

COPD, n (%)							
No	54,029 (93.6)	21,688 (95.2)	11,382 (97.8)	2,770 (96.3)	14,961 (89.8)	1,594 (86.8)	1,634 (83.4)
Yes	3,719 (6.4)	1,100 (4.8)	251 (2.2)	107 (3.7)	1,692 (10.2)	243 (13.2)	326 (16.6)
Diabetes, n (%)							
No	46,563 (80.6)	19,539 (85.7)	10,380 (89.2)	2,489 (86.5)	11,878 (71.3)	1,119 (60.9)	1,158 (59.1)
Yes	11,185 (19.4)	3,249 (14.3)	1,253 (10.8)	388 (13.5)	4,775 (28.7)	718 (39.1)	802 (40.9)
Hypertension, n (%)							
No	37,710 (65.3)	16,023 (70.3)	9,335 (80.2)	2,130 (74.0)	8,643 (51.9)	793 (43.2)	786 (40.1)
Yes	20,038 (34.7)	6,765 (29.7)	2,298 (19.8)	747 (26.0)	8,010 (48.1)	1,044 (56.8)	1,174 (59.9)
Asthma, n (%)							
No	51,726 (89.6)	20,663 (90.7)	10,587 (91.0)	2,517 (87.5)	14,604 (87.7)	1,652 (89.9)	1,703 (86.9)
Yes	6,022 (10.4)	2,125 (9.3)	1,046 (9.0)	360 (12.5)	2,049 (12.3)	185 (10.1)	257 (13.1)
Chronic renal disease, n (%)							

No	53,421 (92.5)	21,585 (94.7)	11,411 (98.1)	2,768 (96.2)	14,520 (87.2)	1,524 (83.0)	1,613 (82.3)
Yes	4,327 (7.5)	1,203 (5.3)	222 (1.9)	109 (3.8)	2,133 (12.8)	313 (17.0)	347 (17.7)
Other chronic respiratory disease, n (%)							
No	55,858 (96.7)	21,900 (96.1)	11,365 (97.7)	2,738 (95.2)	16,186 (97.2)	1,779 (96.8)	1,890 (96.4)
Yes	1,890 (3.3)	888 (3.9)	268 (2.3)	139 (4.8)	467 (2.8)	58 (3.2)	70 (3.6)
Chronic ischemic heart disease, n (%)							
No	52,308 (90.6)	20,978 (92.1)	11,270 (96.9)	2,733 (95.0)	14,236 (85.5)	1,491 (81.2)	1,600 (81.6)
Yes	5,440 (9.4)	1,810 (7.9)	363 (3.1)	144 (5.0)	2,417 (14.5)	346 (18.8)	360 (18.4)
End stage renal disease, n (%)							
No	56,530 (97.9)	22,463 (98.6)	11,560 (99.4)	2,849 (99.0)	16,079 (96.6)	1,740 (94.7)	1,839 (93.8)
Yes	1,218 (2.1)	325 (1.4)	73 (0.6)	28 (1.0)	574 (3.4)	97 (5.3)	121 (6.2)
Liver disease							
No	55,348 (95.8)	21,961 (96.4)	11,407 (98.1)	2,775 (96.5)	15,683 (94.2)	1,718 (93.5)	1,804 (92.0)

Yes	2,400 (4.2)	827 (3.6)	226 (1.9)	102 (3.5)	970 (5.8)	119 (6.5)	156 (8.0)
HIV, n (%)							
No	57,000 (98.7)	22,425 (98.4)	11,533 (99.1)	2,836 (98.6)	16,456 (98.8)	1,811 (98.6)	1,939 (98.9)
Yes	748 (1.3)	363 (1.6)	100 (0.9)	41 (1.4)	197 (1.2)	26 (1.4)	21 (1.1)
Immunocompromised, n (%)							
No	53,530 (92.7)	21,906 (96.1)	11,403 (98.0)	2,802 (97.4)	14,429 (86.6)	1,490 (81.1)	1,500 (76.5)
Yes	4,218 (7.3)	882 (3.9)	230 (2.0)	75 (2.6)	2,224 (13.4)	347 (18.9)	460 (23.5)

AIDS, acquired immunodeficiency syndrome; CCI, Charlson Comorbidity Index; COPD, chronic obstructive pulmonary disease; COVID-19, coronavirus disease 2019; ER, emergency room; HIV, human immunodeficiency virus; ICU, intensive care unit; SD, standard deviation

Supplemental Table 4 Full list of long-term outcomes that occurred >30–≤90 days or >90–≤180 days post COVID-19 diagnosis or hospitalization

	1. Outpatient (N=22,788)		2. ER on diagnosis date (N=11,633)		3. ER (N=2,877)		4. Hospitalization without ICU (N=16,653)		5. ICU without ventilation (N=1,837)		6. ICU with ventilation (N=1,960)	
	>30–≤90 days	**>90–≤180 days**	**>30–≤90 days**	**>90–≤180 days**	**>30–≤90 days**	**>90–≤180 days**	**>30–≤90 days**	**>90–≤180 days**	**>30–≤90 days**	**>90–≤180 days**	**>30–≤90 days**	**>90–≤180 days**
Pneumonia, n (%)												
No	22,368 (98.2)	22,582 (99.1)	11,520 (99.0)	11,579 (99.5)	2,804 (97.5)	2,848 (99.0)	15,612 (93.7)	16,187 (97.2)	1,689 (91.9)	1,768 (96.2)	1,687 (86.1)	1,801 (91.9)
Yes	420 (1.8)	206 (0.9)	113 (1.0)	54 (0.5)	73 (2.5)	29 (1.0)	1,041 (6.3)	466 (2.8)	148 (8.1)	69 (3.8)	273 (13.9)	159 (8.1)
Asthma, n (%)												
No	22,487 (98.7)	22,459 (98.6)	11,532 (99.1)	11,503 (98.9)	2,825 (98.2)	2,822 (98.1)	16,424 (98.6)	16,410 (98.5)	1,810 (98.5)	1,815 (98.8)	1,919 (97.9)	1,922 (98.1)
Yes	301 (1.3)	329 (1.4)	101 (0.9)	130 (1.1)	52 (1.8)	55 (1.9)	229 (1.4)	243 (1.5)	27 (1.5)	22 (1.2)	41 (2.1)	38 (1.9)
COPD, n (%)												

No	22,626 (99.3)	22,776 (99.9)	11,615 (99.8)	11,606 (99.8)	2,865 (99.6)	2,858 (99.3)	16,460 (98.8)	16,442 (98.7)	1,802 (98.1)	1,804 (98.2)	1,894 (96.6)	1,919 (97.9)
Yes	162 (0.7)	179 (0.8)	18 (0.2)	28 (0.2)	12 (0.4)	19 (0.7)	193 (1.2)	211 (1.3)	35 (1.9)	33 (1.8)	66 (3.4)	41 (2.1)
Influenza, n (%)												
No	22,783 (100.0)	22,776 (99.9)	11,630 (100.0)	11,631 (100.0)	2,875 (99.9)	2,877 (100.0)	16,646 (100.0)	16,648 (100.0)	1,837 (100.0)	1,837 (100.0)	1,960 (100.0)	1,960 (100.0)
Yes	5 (0.0)	12 (0.1)	3 (0.0)	2 (0.0)	2 (0.1)	0 (0.0)	7 (0.0)	5 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Stroke, n (%)												
No	22,695 (99.6)	22,696 (99.6)	11,619 (99.9)	11,611 (99.8)	2,865 (99.6)	2,865 (99.6)	16,535 (99.3)	16,506 (99.1)	1,813 (98.7)	1,811 (98.6)	1,935 (98.7)	1,926 (98.3)
Yes	93 (0.4)	92 (0.4)	14 (0.1)	22 (0.2)	12 (0.4)	12 (0.4)	118 (0.7)	147 (0.9)	24 (1.3)	26 (1.4)	25 (1.3)	34 (1.7)
Anxiety, n (%)												
No	22,250 (97.6)	22,169 (97.3)	11,491 (98.8)	11,451 (98.4)	2,779 (96.6)	2,774 (96.4)	16,274 (97.7)	16,268 (97.7)	1,793 (97.6)	1,794 (97.7)	1,889 (96.4)	1,896 (96.7)
Yes	538 (2.4)	619 (2.7)	142 (1.2)	182 (1.6)	98 (3.4)	103 (3.6)	379 (2.3)	385 (2.3)	44 (2.4)	43 (2.3)	71 (3.6)	64 (3.3)

Depression, n (%)												
No	22,456 (98.5)	22,361 (98.1)	11,535 (99.2)	11,502 (98.9)	2,833 (98.5)	2,822 (98.1)	16,375 (98.3)	16,314 (98.0)	1,789 (97.4)	1,786 (97.2)	1,907 (97.3)	1,908 (97.3)
Yes	332 (1.5)	427 (1.9)	98 (0.8)	131 (1.1)	44 (1.5)	55 (1.9)	278 (1.7)	339 (2.0)	48 (2.6)	51 (2.8)	53 (2.7)	52 (2.7)
Myocardial infarction, n (%)												
No	22,691 (99.6)	22,671 (99.5)	11,617 (99.9)	11,617 (99.9)	2,866 (99.6)	2,868 (99.7)	16,497 (99.1)	16,492 (99.0)	1,810 (98.5)	1,806 (98.3)	1,927 (98.3)	1,926 (98.3)
Yes	97 (0.4)	117 (0.5)	16 (0.1)	16 (0.1)	11 (0.4)	9 (0.3)	156 (0.9)	161 (1.0)	27 (1.5)	31 (1.7)	33 (1.7)	34 (1.7)
Interstitial lung disease (fibrosis), n (%)												
No	22,741 (99.8)	22,728 (99.7)	11,623 (99.9)	11,621 (99.9)	2,874 (99.9)	2,873 (99.9)	16,592 (99.6)	16,578 (99.5)	1,830 (99.6)	1,828 (99.5)	1,929 (98.4)	1,922 (98.1)
Yes	47 (0.2)	60 (0.3)	10 (0.1)	12 (0.1)	3 (0.1)	4 (0.1)	61 (0.4)	75 (0.5)	7 (0.4)	9 (0.5)	31 (1.6)	38 (1.9)
Dsypnea, n (%)												
No	21,660 (95.1)	21,567 (94.6)	11,311 (97.2)	11,329 (97.4)	2,675 (93.0)	2,649 (92.1)	15,781 (94.8)	15,838 (95.1)	1,720 (93.6)	1,723 (93.8)	1,759 (89.7)	1,783 (91.0)

Yes	1,128 (4.9)	1,221 (5.4)	322 (2.8)	304 (2.6)	202 (7.0)	228 (7.9)	872 (5.2)	815 (4.9)	117 (6.4)	114 (6.2)	201 (10.3)	177 (9.0)
Respiratory failure, n (%)												
No	22,614 (99.2)	22,654 (99.4)	11,606 (99.8)	11,609 (99.8)	2,863 (99.5)	2,868 (99.7)	16,199 (97.3)	16,380 (98.4)	1,757 (95.6)	1,780 (96.9)	1,685 (86.0)	1,801 (91.9)
Yes	174 (0.8)	134 (0.6)	27 (0.2)	24 (0.2)	14 (0.5)	9 (0.3)	454 (2.7)	273 (1.6)	80 (4.4)	57 (3.1)	275 (14.0)	159 (8.1)
Pulmonary hypertension, n (%)												
No	22,719 (99.7)	22,716 (99.7)	11,626 (99.9)	11,622 (99.9)	2,872 (99.8)	2,874 (99.9)	16,566 (99.5)	16,551 (99.4)	1,824 (99.3)	1,821 (99.1)	1,944 (99.2)	1,927 (98.3)
Yes	69 (0.3)	72 (0.3)	7 (0.1)	11 (0.1)	5 (0.2)	3 (0.1)	87 (0.5)	102 (0.6)	13 (0.7)	16 (0.9)	16 (0.8)	33 (1.7)
Pulmonary embolism, n (%)												
No	22,714 (99.7)	22,719 (99.7)	11,622 (99.9)	11,623 (99.9)	2,865 (99.6)	2,863 (99.5)	16,478 (98.9)	16,513 (99.2)	1,809 (98.5)	1,815 (98.8)	1,918 (97.9)	1,938 (98.9)
Yes	74 (0.3)	69 (0.3)	11 (0.1)	10 (0.1)	12 (0.4)	14 (0.5)	175 (1.1)	140 (0.8)	28 (1.5)	22 (1.2)	42 (2.1)	22 (1.1)
Bronchitis, n (%)												

No	22,707 (99.6)	22,705 (99.6)	11,611 (99.8)	11,619 (99.9)	2,862 (99.5)	2,868 (99.7)	16,583 (99.6)	16,598 (99.7)	1,830 (99.6)	1,832 (99.7)	1,939 (98.9)	1,940 (99.0)
Yes	81 (0.4)	83 (0.4)	22 (0.2)	14 (0.1)	15 (0.5)	9 (0.3)	70 (0.4)	55 (0.3)	7 (0.4)	5 (0.3)	21 (1.1)	20 (1.0)
Emphysema, n (%)												
No	22,727 (99.7)	22,722 (99.7)	11,626 (99.9)	11,620 (99.9)	2,872 (99.8)	2,870 (99.8)	16,591 (99.6)	16,577 (99.5)	1,815 (98.8)	1,822 (99.2)	1,941 (99.0)	1,944 (99.2)
Yes	61 (0.3)	66 (0.3)	7 (0.1)	13 (0.1)	5 (0.2)	7 (0.2)	62 (0.4)	76 (0.5)	22 (1.2)	15 (0.8)	19 (1.0)	16 (0.8)
Bronchiectasis, n (%)												
No	22,765 (99.9)	22,763 (99.9)	11,632 (100.0)	11,629 (100.0)	2,876 (100.0)	2,874 (99.9)	16,630 (99.9)	16,625 (99.8)	1,836 (99.9)	1,836 (99.9)	1,951 (99.5)	1,952 (99.6)
Yes	23 (0.1)	25 (0.1)	1 (0.0)	4 (0.0)	1 (0.0)	3 (0.1)	23 (0.1)	28 (0.2)	1 (0.1)	1 (0.1)	9 (0.5)	8 (0.4)
Encephalopathy, n (%)												
No	22,709 (99.7)	22,732 (99.8)	11,624 (99.9)	11,627 (99.9)	2,872 (99.8)	2,874 (99.9)	16,545 (99.4)	16,554 (99.4)	1,809 (98.5)	1,816 (98.9)	1,911 (97.5)	1,923 (98.1)
Yes	79 (0.3)	56 (0.2)	9 (0.1)	6 (0.1)	5 (0.2)	3 (0.1)	108 (0.6)	99 (0.6)	28 (1.5)	21 (1.1)	49 (2.5)	37 (1.9)

Memory loss, n (%)												
No	22,752 (99.8)	22,716 (99.7)	11,622 (99.9)	11,621 (99.9)	2,872 (99.8)	2,870 (99.8)	16,626 (99.8)	16,599 (99.7)	1,833 (99.8)	1,831 (99.7)	1,951 (99.5)	1,946 (99.3)
Yes	36 (0.2)	72 (0.3)	11 (0.1)	12 (0.1)	5 (0.2)	7 (0.2)	27 (0.2)	54 (0.3)	4 (0.2)	6 (0.3)	9 (0.5)	14 (0.7)
Confusion or disorientation, n (%)												
No	22,699 (99.6)	22,706 (99.6)	11,621 (99.9)	11,617 (99.9)	2,869 (99.7)	2,869 (99.7)	16,531 (99.3)	16,526 (99.2)	1,817 (98.9)	1,817 (98.9)	1,929 (98.4)	1,939 (98.9)
Yes	89 (0.4)	82 (0.4)	12 (0.1)	16 (0.1)	8 (0.3)	8 (0.3)	122 (0.7)	127 (0.8)	20 (1.1)	20 (1.1)	31 (1.6)	21 (1.1)
Dementia, n (%)												
No	22,694 (99.6)	22,709 (99.7)	11,628 (100.0)	11,625 (99.9)	2,870 (99.8)	2,872 (99.8)	16,494 (99.0)	16,494 (99.0)	1,810 (98.5)	1,816 (98.9)	1,944 (99.2)	1,947 (99.3)
Yes	94 (0.4)	79 (0.3)	5 (0.0)	8 (0.1)	7 (0.2)	5 (0.2)	159 (1.0)	159 (1.0)	27 (1.5)	21 (1.1)	16 (0.8)	13 (0.7)
Cardiac arrhythmia, n (%)												
No	22,627 (99.3)	22,598 (99.2)	11,594 (99.7)	11,593 (99.7)	2,860 (99.4)	2,850 (99.1)	16,515 (99.2)	16,488 (99.0)	1,819 (99.0)	1,816 (98.9)	1,935 (98.7)	1,931 (98.5)

Yes	161 (0.7)	190 (0.8)	39 (0.3)	40 (0.3)	17 (0.6)	27 (0.9)	138 (0.8)	165 (1.0)	18 (1.0)	21 (1.1)	25 (1.3)	29 (1.5)
Respiratory, n (%)												
No	20,942 (91.9)	20,942 (91.9)	11,113 (95.5)	11,149 (95.8)	2,555 (88.8)	2,566 (89.2)	14,529 (87.2)	15,054 (90.4)	1,541 (83.9)	1,615 (87.9)	1,413 (72.1)	1,579 (80.6)
Yes	1,846 (8.1)	1,846 (8.1)	520 (4.5)	484 (4.2)	322 (11.2)	311 (10.8)	2,124 (12.8)	1,599 (9.6)	296 (16.1)	222 (12.1)	547 (27.9)	381 (19.4)
CV, n (%)												
No	22,349 (98.1)	22,313 (97.9)	11,522 (99.3)	11,541 (99.2)	2,823 (98.1)	2,820 (98.0)	16,068 (96.5)	16,035 (96.3)	1,735 (94.4)	1,738 (94.6)	1,832 (93.5)	1,828 (93.3)
Yes	439 (1.9)	475 (2.1)	81 (0.7)	92 (0.8)	54 (1.9)	57 (2.0)	585 (3.5)	618 (3.7)	102 (5.6)	99 (5.4)	128 (6.5)	132 (6.7)
Mental health, n (%)												
No	21,848 (95.9)	21,707 (95.3)	11,402 (98.0)	11,340 (97.5)	2,733 (95.0)	2,723 (94.6)	15,834 (95.1)	15,787 (94.8)	1,718 (93.5)	1,719 (93.6)	1,793 (91.5)	1,816 (92.7)
Yes	940 (4.1)	1,081 (4.7)	231 (2.0)	293 (2.5)	144 (5.0)	154 (5.4)	819 (4.9)	866 (5.2)	119 (6.5)	118 (6.4)	167 (8.5)	144 (7.3)
Cancer, n (%)												

No	22,561 (99.4)	22,637 (99.3)	11,609 (99.8)	11,603 (99.7)	2,859 (99.4)	2,863 (99.5)	16,558 (99.4)	16,547 (99.4)	1,828 (99.5)	1,825 (99.3)	1,938 (98.9)	1,946 (99.3)
Yes	137 (0.6)	151 (0.7)	24 (0.2)	30 (0.3)	18 (0.6)	14 (0.5)	95 (0.6)	106 (0.6)	9 (0.5)	12 (0.7)	22 (1.1)	14 (0.7)
Respiratory and CV, n (%)												
No	22,657 (99.4)	22,664 (99.5)	11,617 (99.9)	11,603 (99.7)	2,860 (99.4)	2,858 (99.3)	16,452 (98.8)	16,472 (98.9)	1,807 (98.8)	1,811 (98.6)	1,897 (96.8)	1,913 (97.6)
Yes	131 (0.6)	124 (0.5)	16 (0.1)	30 (0.3)	17 (0.6)	19 (0.7)	201 (1.2)	181 (1.1)	30 (1.6)	26 (1.4)	63 (3.2)	47 (2.4)
Respiratory and mental health, n (%)												
No	22,583 (99.1)	22,572 (99.1)	11,568 (99.4)	11,570 (99.5)	2,829 (98.3)	2,844 (98.9)	16,391 (98.4)	16,408 (98.5)	1,792 (97.6)	1,803 (98.1)	1,877 (95.8)	1,910 (97.4)
Yes	205 (0.9)	216 (0.9)	65 (0.6)	63 (0.5)	48 (1.7)	33 (1.1)	262 (1.6)	245 (1.5)	45 (2.4)	34 (1.9)	83 (4.2)	50 (2.6)
Mental health and CV, n (%)												
No	22,735 (99.8)	22,739 (99.8)	11,625 (99.9)	11,624 (99.9)	2,873 (99.9)	2,875 (99.9)	16,596 (99.7)	16,589 (99.6)	1,827 (99.5)	1,826 (99.4)	1,953 (99.6)	1,949 (99.4)
Yes	53 (0.2)	49 (0.2)	8 (0.1)	9 (0.1)	4 (0.1)	2 (0.1)	57 (0.3)	64 (0.4)	10 (0.5)	11 (0.6)	7 (0.4)	11 (0.6)

Respiratory, CV, and mental health, n (%)												
No	22,731 (99.7)	22,736 (99.8)	11,624 (99.9)	11,623 (99.9)	2,871 (99.8)	2,868 (99.7)	16,555 (99.4)	16,569 (99.5)	1,820 (99.1)	1,822 (99.2)	1,930 (98.5)	1,929 (98.4)
Yes	57 (0.3)	52 (0.2)	9 (0.1)	10 (0.1)	6 (0.2)	9 (0.3)	98 (0.6)	84 (0.5)	17 (0.9)	15 (0.8)	30 (1.5)	31 (1.6)
No conditions, n (%)												
No	2,722 (11.9)	2,909 (12.8)	725 (6.2)	747 (6.4)	439 (15.3)	450 (15.6)	2,812 (16.9)	2,425 (14.6)	398 (21.7)	338 (18.4)	629 (32.1)	487 (24.8)
Yes	20,066 (88.1)	19,879 (87.2)	10,908 (93.8)	10,886 (93.6)	2,438 (84.7)	2,427 (84.4)	13,841 (83.1)	14,228 (85.4)	1,439 (78.3)	1,499 (81.6)	1,331 (67.9)	1,473 (75.2)

COPD, chronic obstructive pulmonary disease; COVID-19, coronavirus disease 2019; CV, cardiovascular; ER, emergency room; ICU, intensive care unit

Supplemental Table 5 Full list of risk ratios and 95% confidence intervals of covariates associated with the occurrence of new respiratory, cardiovascular, and mental health conditions at >30–≤90 days or >90–≤180 days post COVID-19 diagnosis or hospital discharge

RR (95% CI)	Respiratory conditions		Cardiovascular conditions		Mental health conditions	
	>30–≤90 days	>90–≤180 days	>30–≤90 days	>90–≤180 days	>30–≤90 days	>90–≤180 days
Age group						
18–29 years	1.18 (0.80, 1.79)	1.08 (0.76, 1.56)	NA*	3.18 (0.94, 19.53)	2.67 (1.59, 4.85)	1.81 (1.22, 2.79)
30–39 years	1.90 (1.32, 2.82)	1.59 (1.14, 2.27)	NA*	6.09 (1.88, 36.34)	3.00 (1.79, 5.43)	1.87 (1.26, 2.88)
40–49 years	2.44 (1.71, 3.58)	1.96 (1.41, 2.77)	NA*	6.68 (2.05, 39.85)	2.73 (1.63, 4.96)	1.91 (1.29, 2.94)
50–64 years	2.84 (2.01, 4.13)	1.92 (1.40, 2.71)	NA*	8.03 (2.52, 47.28)	2.94 (1.77, 5.28)	1.63 (1.11, 2.50)
65–74 years	2.68 (1.86, 3.95)	1.88 (1.34, 2.70)	NA*	10.46 (3.10, 62.22)	2.30 (1.35, 4.22)	1.28 (0.85, 2.02)
75–84 years	2.43 (1.65, 3.65)	1.52 (1.04, 2.24)	NA*	18.40 (5.23, 107.58)	2.61 (1.50, 4.85)	1.29 (0.83, 2.07)
≥85 years	1.80 (1.10, 2.95)	1.62 (1.01, 2.58)	NA*	15.49 (2.47, 111.25)	3.05 (1.59, 6.02)	1.73 (1.00, 2.98)
Sex						
Male	0.91 (0.84, 0.99)	0.84 (0.77, 0.92)	1.18 (0.92, 1.51)	1.26 (0.98, 1.61)	0.63 (0.56, 0.70)	0.53 (0.47, 0.59)
Race						

Caucasian	1.02 (0.94, 1.12)	1.13 (1.02, 1.24)	1.10 (0.83, 1.47)	1.10 (0.83, 1.48)	1.48 (1.31, 1.69)	1.792 (1.59, 2.03)
Asian	0.99 (0.81, 1.19)	1.00 (0.80, 1.24)	0.72 (0.33, 1.36)	0.84 (0.41, 1.56)	0.74 (0.50, 1.04)	0.81 (0.56, 1.12)
Ethnicity						
Non-Hispanic	1.12 (0.97, 1.29)	1.38 (1.17, 1.62)	2.56 (1.53, 4.63)	2.12 (1.30, 3.70)	1.40 (1.14, 1.73)	1.54 (1.27, 1.87)
Obesity	1.33 (1.21, 1.46)	1.38 (1.24, 1.53)	1.08 (0.76, 1.51)	1.18 (0.83, 1.63)	1.41 (1.24, 1.59)	1.49 (1.32, 1.67)
Insurance						
Medicaid	1.01 (0.89, 1.15)	1.07 (0.94, 1.22)	1.24 (0.85, 1.76)	1.23 (0.84, 1.75)	1.39 (1.18, 1.64)	1.41 (1.21, 1.63)
Medicare	1.06 (0.93, 1.21)	1.06 (0.91, 1.22)	1.45 (0.94, 2.21)	1.51 (1.00, 2.26)	1.53 (1.29, 1.82)	1.46 (1.23, 1.72)
Other payor type	0.77 (0.64, 0.93)	1.07 (0.90, 1.27)	1.00 (0.56, 1.65)	1.10 (0.64, 1.77)	0.97 (0.75, 1.23)	1.03 (0.82, 1.28)
Uninsured	0.71 (0.58, 0.85)	0.65 (0.52, 0.80)	0.96 (0.55, 1.57)	0.66 (0.33, 1.16)	0.76 (0.56, 0.99)	0.63 (0.47, 0.82)
Sub-cohort						
ER on diagnosis	0.64 (0.56, 0.74)	0.56 (0.48, 0.65)	0.45 (0.27, 0.71)	0.59 (0.38, 0.89)	0.60 (0.49, 0.72)	0.65 (0.55, 0.78)
ER	1.39 (1.17, 1.65)	1.33 (1.10, 1.58)	1.41 (0.83, 2.27)	1.57 (0.95, 2.47)	1.22 (0.95, 1.54)	1.20 (0.95, 1.50)
Hospitalisation without ICU	1.33 (1.20, 1.47)	1.02 (0.91, 1.14)	1.74 (1.28, 2.35)	1.42 (1.04, 1.94)	1.03 (0.89, 1.18)	1.08 (0.95, 1.23)
ICU without ventilation	1.69 (1.39, 2.03)	1.18 (0.93, 1.47)	1.69 (0.75, 3.28)	2.41 (1.25, 4.23)	1.31 (0.99, 1.71)	1.34 (1.02, 1.73)

ICU with ventilation	2.64 (2.27, 3.04)	1.86 (1.55, 2.21)	3.16 (1.83, 5.18)	2.65 (1.49, 4.43)	1.89 (1.51, 2.35)	1.52 (1.20, 1.91)
Month of COVID-19 diagnosis						
Feb–Apr 2020	1.08 (0.97, 1.20)	1.06 (0.93, 1.19)	0.88 (0.62, 1.24)	0.96 (0.66, 1.38)	0.85 (0.72, 1.01)	0.93 (0.80, 1.09)
May 2020	0.75 (0.67, 0.83)	0.92 (0.82, 1.04)	0.76 (0.55, 1.06)	1.21 (0.89, 1.65)	0.82 (0.71, 0.95)	0.90 (0.79, 1.03)
Jun 2020	0.58 (0.51, 0.65)	0.72 (0.63, 0.81)	0.68 (0.48, 0.96)	0.75 (0.52, 1.07)	0.72 (0.62, 0.84)	0.74 (0.64, 0.86)
Jul 2020	0.48 (0.34, 0.65)	0.61 (0.45, 0.82)	0.33 (0.08, 0.88)	0.81 (0.34, 1.65)	0.60 (0.40, 0.87)	0.79 (0.57, 1.08)
Weighted CCI	1.05 (1.02, 1.07)	1.07 (1.04, 1.09)	1.17 (1.10, 1.25)	1.16 (1.08, 1.23)	1.12 (1.09, 1.15)	1.14 (1.11, 1.17)

Grey and blue shading denote increased and decreased risk of a new condition occurring, respectively. Reference groups are: <18 years (age), female (sex), African American (race), Hispanic (ethnicity), non-obese (obesity), diagnosis in February and March 2020 (diagnosis month); commercial (insurance), outpatient (sub-cohort).

*Values were not calculable as the reference group (<18 years) had no new diagnoses of clinical conditions

CCI, Charlson Comorbidity Index; CI, confidence interval; COVID-19, coronavirus disease 2019; ER, emergency room; ICU, intensive care unit;

NA, not applicable; RR, risk ratio

Supplemental Table 6 Full list of risk ratios and 95% confidence intervals of covariates associated with a new cancer diagnosis at >30–≤90 days or >90–≤180 days post COVID-19 diagnosis or hospital discharge

	Risk ratio (95% CI)	
	>30–≤90 days	**>90–≤180 days**
Age group		
18–29 years	0.95 (0.31, 4.16)	1.17 (0.19, 22.13)
30–39 years	0.89 (0.29, 3.90)	2.63 (0.51, 46.94)
40–49 years	1.43 (0.49, 6.05)	3.38 (0.69, 59.46)
50–64 years	2.44 (0.90, 9.93)	8.35 (1.84, 138.73)
65–74 years	3.19 (1.13, 13.21)	13.50 (2.91, 217.30)
75–84 years	2.60 (0.86, 11.12)	15.50 (3.25, 247.65)
≥85 years	2.12 (0.53, 10.33)	8.45 (1.39, 151.23)
Sex		
Male	0.97 (0.75, 1.27)	1.03 (0.79, 1.33)
Race		
Caucasian	1.51 (1.10, 2.11)	1.05 (0.78, 1.43)
Asian	1.45 (0.66, 2.80)	1.15 (0.53, 2.20)
Ethnicity		
Non-Hispanic	2.49 (1.33, 5.28)	1.34 (0.79, 2.50)

Obesity	1.34 (0.98, 1.80)	1.34 (1.00, 1.79)
Insurance		
Medicaid	0.88 (0.52, 1.41)	0.68 (0.38, 1.13)
Medicare	1.43 (0.99, 2.07)	1.10 (0.78, 1.55)
Other payor type	0.72 (0.32, 1.39)	0.69 (0.32, 1.29)
Uninsured	1.32 (0.68, 2.32)	0.49 (0.17, 1.09)
Sub-cohort		
ER on diagnosis	0.47 (0.27, 0.76)	0.53 (0.32, 0.83)
ER	1.23 (0.68, 2.06)	0.69 (0.32, 1.30)
Hospitalization without ICU	0.71 (0.50, 0.99)	0.65 (0.47, 0.90)
ICU without ventilation	0.45 (0.16, 1.01)	0.50 (0.21, 1.02)
ICU with ventilation	1.24 (0.71, 2.06)	0.56 (0.27, 1.04)
Month of COVID-19 diagnosis		
Feb–Apr 2020	1.44 (0.98, 2.11)	1.72 (1.17, 2.51)
May 2020	1.21 (0.85, 1.72)	1.18 (0.82, 1.69)
Jun 2020	0.90 (0.61, 1.32)	1.28 (0.89, 1.84)
Jul 2020	0.70 (0.21, 1.71)	0.80 (0.24, 1.96)
Weighted CCI	1.04 (0.97, 1.11)	1.06 (0.99, 1.12)

Reference groups are: <18 years (age), female (sex), African American (race), Hispanic (ethnicity), non-obese (obesity); diagnosis in February and March 2020 (diagnosis month), commercial (insurance), outpatient (sub-cohort).

CCI, Charlson Comorbidity Index; CI, confidence interval;
COVID-19, coronavirus disease 2019; ER, emergency room; ICU, intensive care unit

The RECORD statement – checklist of items, extended from the STROBE statement, that should be reported in observational studies using routinely collected health data.

	Item No.	STROBE items	Location in manuscript where items are reported	RECORD items	Location in manuscript where items are reported
Title and abstract					
	1	(a) Indicate the study's design with a commonly used term in the title or the abstract (b) Provide in the abstract an informative and balanced summary of what was done and what was found	*(a) confirmed (Design section) (b) confirmed adequately covered in abstract*	RECORD 1.1: The type of data used should be specified in the title or abstract. When possible, the name of the databases used should be included. RECORD 1.2: If applicable, the geographic region and timeframe within which the study took place should be reported in the title or abstract. RECORD 1.3: If linkage between databases was conducted for the study, this should be clearly stated in the title or abstract.	*Abstract (Objective and Design)* *Abstract (Setting and Participants)* *Not applicable*
Introduction					
Background rationale	2	Explain the scientific background and rationale for the investigation being reported	*Confirmed present in introduction*		
Objectives	3	State specific objectives, including any prespecified hypotheses	*Specific objective stated (last paragraph of introduction; there were no pre-*		

			specified hypotheses)		
Methods					
Study Design	4	Present key elements of study design early in the paper	*Included in methods ('Patients and study design') and described in Figure 1*		
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	*Included in methods ('Database' & 'Patients and study design' sections)*		

Participants	6	*(a) Cohort study*- Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*- Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*- Give the eligibility criteria, and the sources and methods of selection of participants *(b) Cohort study*- For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*- For matched studies, give matching criteria and the number of controls per case	*(a) confirmed included in methods ('Patients and study design' section (b) not relevant (not a matched study)*	RECORD 6.1: The methods of study population selection (such as codes or algorithms used to identify subjects) should be listed in detail. If this is not possible, an explanation should be provided. RECORD 6.2: Any validation studies of the codes or algorithms used to select the population should be referenced. If validation was conducted for this study and not published elsewhere, detailed methods and results should be provided. RECORD 6.3: If the study involved linkage of databases, consider use of a flow diagram or other graphical display to demonstrate the data linkage process, including the number of individuals with linked data at each stage.	*Confirmed in methods ('Patients and study design')* *The algorithms have been used previously and is cited in the methods (Chawla et al., 2021)* *Not applicable*
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable.	*All definitions are presented in the methods ('Patients and study design', 'Modelling and*	RECORD 7.1: A complete list of codes and algorithms used to classify exposures, outcomes, confounders, and effect modifiers should be provided. If these cannot be reported, an explanation should be provided.	*Relevant lists are provided throughout the manuscript (e.g. ICD-10 codes in*

			statistical analysis', and 'Sensitivity analysis' sections)		*supplemental Tables 1 and 2, list of confounders in methods section 'Modeling and statistical analysis')*
Data sources/ measurement	8	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	*Source of data is the Optum Electronic Medical Record data, and are routinely collected by practicing physicians (detailed in methods section)*		
Bias	9	Describe any efforts to address potential sources of bias	*A sensitivity analysis was performed, and relevant controls (non-hospital setting covariates) were included in our statistical models*		
Study size	10	Explain how the study size was arrived at	*All eligible patients in the Optum dataset were included, without a prespecified study size (explained in the*		

			database and patients and study design section)		
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why	*Described in Methods section 'Modeling and statistical analysis'*		
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding (b) Describe any methods used to examine subgroups and interactions (c) Explain how missing data were addressed (d) *Cohort study*- If applicable, explain how loss to follow-up was addressed *Case-control study*- If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*- If applicable, describe analytical methods taking account of sampling strategy (e) Describe any sensitivity analyses	*a) Methods ('Modeling and statistical analysis')* *b) We do not conduct sub-group analysis* *c) Explained in discussion section* *d) We have conducted a retrospective cohort study. Regarding the loss-to follow-up, since we are not assessing the effect of a treatment, rather looking at disease severity, we assume it is non-differential.* *e) as described in methods ('Sensitivity analysis')*		

Data access and cleaning methods		..		RECORD 12.1: Authors should describe the extent to which the investigators had access to the database population used to create the study population.	*Authors had access to deidentified EMR data*
				RECORD 12.2: Authors should provide information on the data cleaning methods used in the study.	*Data cleaning methods have been described previously; the reference is cited in the Methods 'Database' section (Chawla et al).*
Linkage		..		RECORD 12.3: State whether the study included person-level, institutional-level, or other data linkage across two or more databases. The methods of linkage and methods of linkage quality evaluation should be provided.	*EMR data from hospital networks were used to form the Optum dataset. Linkage of EMR data and methods are described on Optum's website: https://www.optum.com/business/solutions/life-sciences/real-world-data/ehr-data.html*

Results					
Participants	13	(a) Report the numbers of individuals at each stage of the study (*e.g.*, numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed) (b) Give reasons for non-participation at each stage. (c) Consider use of a flow diagram	*The number of patients in the dataset and those with a COVID-19 diagnosis is given in the Methods 'Database' section. The criteria on how the population is selected is made clear in methods 'Patients and study design' section (inclusion and exclusion criteria).*	RECORD 13.1: Describe in detail the selection of the persons included in the study (*i.e.,* study population selection) including filtering based on data quality, data availability and linkage. The selection of included persons can be described in the text and/or by means of the study flow diagram.	*The criteria on how the population is selected is made clear in methods 'Patients and study design' section (inclusion and exclusion criteria)*
Descriptive data	14	(a) Give characteristics of study participants (*e.g.*, demographic, clinical, social) and information on exposures and potential confounders (b) Indicate the number of participants with missing data for each variable of interest (c) *Cohort study*- summarise follow-up time (*e.g.*, average and total amount)	*Shown in detail in tables e.g. Table 1 and also in appendix (table 1 and 2)*		
Outcome data	15	*Cohort study*- Report numbers of outcome events or summary measures over time *Case-control study*-	*Outcome data are presented in Table 2*		

		Report numbers in each exposure			
		category, or summary measures of exposure *Cross-sectional study*- Report numbers of outcome events or summary measures			
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder- adjusted estimates and their precision (e.g., 95% confidence interval). Make clear which confounders were adjusted for and why they were included (b) Report category boundaries when continuous variables were categorized (c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	*Confounder-adjusted estimates are provided; however, unadjusted results could be derived from Table 2, where the raw counts and percentages can be used to calculate raw measures of effect. Confounders we control for are described in the modelling and statistical analysis section*		
Other analyses	17	Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses	*Sensitivity analysis is reported*		

Discussion					
Key results	18	Summarise key results with reference to study objectives	*Covered in discussion*		
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	*An extensive limitations section is included, covering the relevant aspects*	RECORD 19.1: Discuss the implications of using data that were not created or collected to answer the specific research question(s). Include discussion of misclassification bias, unmeasured confounding, missing data, and changing eligibility over time, as they pertain to the study being reported.	*An extensive limitations section is included, covering the relevant aspects*
Interpretation	20	Give a cautious overall interpretation of results considering objectives,	*Covered in discussion*		
		limitations, multiplicity of analyses, results from similar studies, and other relevant evidence			
Generalisability	21	Discuss the generalisability (external validity) of the study results	*Covered in discussion*		
Other Information					
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present	*Covered in funding section*		

		article is based			
Accessibility of protocol, raw data, and programming code		..		RECORD 22.1: Authors should provide information on how to access any supplemental information such as the study protocol, raw data, or programming code.	*Information is included in the data availability statement*

*Reference: Benchimol EI, Smeeth L, Guttmann A, Harron K, Moher D, Petersen I, Sørensen HT, von Elm E, Langan SM, the RECORD Working Committee. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. *PLoS Medicine* 2015; in press.

BMJ Open

Severity of COVID-19 and adverse long-term outcomes: a retrospective cohort study based on a US electronic health record database

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-056284.R1
Article Type:	Original research
Date Submitted by the Author:	04-Nov-2021
Complete List of Authors:	Jovanoski, Nick; F Hoffmann-La Roche Ltd Chen, Xin; Genentech Inc Becker, Ursula; F Hoffmann-La Roche Ltd Zalocusky, Kelly; Genentech Inc Chawla, Devika; Genentech Inc Tsai, Larry; Genentech Inc Borm, Michelle; F Hoffmann-La Roche Ltd Neighbors, Margaret; Genentech Inc Yau, Vincent; Genentech Inc
Primary Subject Heading:	Respiratory medicine
Secondary Subject Heading:	Epidemiology
Keywords:	COVID-19, Epidemiology < INFECTIOUS DISEASES, PUBLIC HEALTH, Respiratory infections < THORACIC MEDICINE

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Severity of COVID-19 and adverse long-term outcomes: a retrospective cohort study based on a US electronic health record database

Nick Jovanoski[1], Xin Chen[2], Ursula Becker[1], Kelly Zalocusky[2], Devika Chawla[2], Larry Tsai[2], Michelle Borm[3], Margaret Neighbors[2], Vincent Yau[2]

1. F. Hoffmann-La Roche Ltd, Basel, Switzerland

2. Genentech, Inc., South San Francisco, CA, USA

3. Roche Nederland BV, Woerden, Netherlands.

Author for correspondence: Vincent Yau, Ph.D. Genentech, Inc., South San Francisco, CA, USA. yau.vincent@gene.com

Target journal: BMJ Open

Word count: 3694/4000

Figures / Tables: 4 / 2 (plus 1 / 6 supplementary)

References: 28

Abstract (300/300 words)

Objective: To identify potential risk factors for adverse long-term outcomes (LTOs) associated with coronavirus disease 2019 (COVID-19), using a large electronic health record (EHR) database.

Design: Retrospective cohort study. Patients with COVID-19 were assigned into sub-cohorts according to the most intensive treatment setting experienced. Newly diagnosed conditions were classified as respiratory, cardiovascular, or mental health LTOs at either >30–≤90 days or >90–≤180 days after COVID-19 diagnosis or hospital discharge. Multivariate regression analysis was performed to identify any effect of disease severity on LTO incidence.

Setting: Optum® de-identified COVID-19 EHR dataset drawn from hospitals and clinics across the United States.

Participants: Individuals diagnosed with COVID-19 (N=57,748) between February 20, 2020 and July 4, 2020.

Main outcomes: Incidence of new clinical conditions after COVID-19 diagnosis or hospital discharge and the potential effect of disease severity on their risk of occurrence.

Results: Patients were assigned into one of six sub-cohorts: outpatient (n=22,788), emergency room (ER) with same-day COVID-19 diagnosis (n=11,633), ER with COVID-19 diagnosis ≤21 days before ER visit (n=2,877), hospitalization without intensive care unit (ICU; n=16,653), ICU without ventilation (n=1,837), and ICU with ventilation (n=1,960). Respiratory LTOs were more common than cardiovascular or mental health LTOs across sub-cohorts, and LTO incidence was higher in hospitalized versus non-hospitalized sub-cohorts. Patients with the most severe disease were at increased risk of respiratory (risk ratio [RR] 1.86, 95% confidence interval [CI] 1.56, 2.21), cardiovascular (RR 2.65, 95% CI 1.49, 4.43), and mental health outcomes (RR 1.52, 95% CI 1.20, 1.91) up to six months after hospital discharge compared with outpatients.

Conclusions: Patients with severe COVID-19 had increased risk of new clinical conditions up to six months after hospital discharge. The extent that treatment setting (e.g., ICU) contributed to these conditions is unknown, but strategies to prevent COVID-19 progression may nonetheless minimise their occurrence.

Strengths and limitations of this study

- This study used a large electronic health record database containing a rich source of patient-level medical and administrative records from hospitals, emergency departments, and outpatient centers across the United States.
- Multivariate logistic regression analysis was used to adjust for measured confounders and assess the effect of increasing COVID-19 severity (proxied by treatment setting) on the risk of new clinical conditions being diagnosed up to six months after COVID-19 diagnosis or hospital discharge.
- A sensitivity analysis assessing the effect of increasing COVID-19 severity on the risk of a new cancer diagnosis served as a negative control.
- The main limitation of this retrospective study is that we use treatment setting as a proxy for COVID-19 severity, and therefore it is difficult to tease out effects specific to the treatment setting (e.g., invasive ventilation) from the underlying COVID-19 severity; any differences that exist between cohorts could bias the results, and as all potential confounders may not be controlled for, the results do not indicate causality.
- Additional limitations include missing information on smoking status, the lack of a COVID-19-negative control group, the possibility of missing data, being restricted to examining conditions captured by ICD-10 codes, the lack of information on COVID-19 treatments received, and the lack of laboratory values or other biomarkers to better characterize disease.

Introduction

The coronavirus disease 2019 (COVID-19) pandemic caused by the novel severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) has imposed an immense burden of morbidity and mortality worldwide.[1] Although the majority of patients experience mild or moderate symptoms that resolve within a few weeks of initial infection, increasing evidence suggests that a subset of patients continue to display symptoms beyond four weeks after infection.[2 3] These symptoms are wide ranging and often extend beyond the typical initial symptoms of COVID-19 to include respiratory (e.g., dyspnea, decreased exercise capacity), cardiovascular (e.g., heart palpitations, chest pain), and mental health (e.g., confusion, disorientation) disorders.[4 5] Notably, such outcomes have been observed even in patients with mild acute COVID-19 symptoms.[6] These prolonged symptoms have collectively been referred to by several names including post-acute COVID-19 (PAC), post-COVID-19 syndrome (PCS), post-acute sequelae of SARS-CoV-2 infection (PASC), and possibly more commonly 'long COVID'.[7 8] However, due to the overlapping and non-specific range of symptoms experienced, the medical community has not yet converged on precise definitions, and it is possible that distinct subsets of long COVID patients exist. It has also been suggested that long COVID can be further sub-divided into subacute COVID-19 (4–12 weeks after initial onset of COVID-19 symptoms) and post-COVID-19 syndrome (beyond 12 weeks).[4 9] The underlying pathogenic mechanisms of long COVID are not well understood, but multiple causes have been proposed, including immune dysregulation and viral persistence.[10] Additionally, in patients with severe disease requiring treatment in the intensive care unit (ICU), non-specific secondary effects cannot be ruled out, similar to those observed in 'post-intensive care syndrome'.[11]

High-quality clinical data on respiratory, cardiovascular, and neurologic sequelae of SARS-CoV-2 infection are beginning to emerge,[12-14] and several observational studies and patient registries have been established to better understand the long-term outcomes (LTOs)

of COVID-19.[15 16] However, little is known about the potential baseline factors that may predict the development of long COVID.

Retrospective cohort studies using electronic health records (EHRs) are uniquely positioned due to their size and convenience to provide insights into factors underlying long COVID development and the range of long COVID conditions that exist. The Optum® de-identified COVID-19 EHR dataset contains patient-level medical and administrative records from hospitals, emergency departments, outpatient centers, and laboratories across the United States (US). This dataset has previously been utilized to describe key epidemiological features of a large cohort of hospitalized patients with COVID-19[17] and to develop a prognostic model of in-hospital mortality.[18]

The current study utilized the Optum® de-identified COVID-19 EHR dataset to better understand the types of LTOs encountered by patients with long COVID, to define the factors that predict their diagnosis, and to understand the role COVID-19 severity plays in the manifestation of these outcomes.

Methods

Database

Individuals with COVID-19 diagnosed between February 20, 2020 and July 4, 2020 were extracted from the Optum® de-identified COVID-19 EHR dataset (569,149 individuals from 3,832,315 in the entire dataset). This dataset contains patient-level medical and administrative records from hospitals, emergency departments, outpatient centers, and laboratories across the US. All data were de-identified according to the Health Insurance Portability and Accountability Act Expert Method and managed according to Optum® customer data use agreements. The COVID-19 EHR dataset comprises clinical information sourced from hospital networks that provided data meeting Optum®'s internal data quality criteria. Data cleaning methods used were as described previously.[17]

Patients and study design

Eligible patients (overall COVID-19 cohort) had ≥1 of the following: a COVID-19 diagnosis code (U07.1, U07.2), a positive diagnostic test for SARS-CoV-2 infection (e.g., molecular or antigen test), or a B97.29 diagnosis code (other coronavirus as the cause of diseases classified elsewhere) without a negative SARS-CoV-2 molecular test within 14 days. The index date was defined as the date of COVID-19 diagnosis or COVID-19-related hospitalization (as defined below), whichever occurred first. The baseline period was defined as the 12 months prior to the index date, and a minimum of 180 days follow-up was required for all patients. The overall study design is shown in **Figure 1**.

Eligible patients were assigned into the following six sub-cohorts according to treatment setting: **1. Outpatient**, patients with a COVID-19 diagnosis and no record of hospitalization or an emergency room (ER) visit within 21 days of diagnosis; **2. ER on diagnosis**, COVID-19 diagnosis on the same day as ER visit; **3. ER**, COVID-19 diagnosis prior to ER visit, i.e., patients with an ER visit within 21 days after COVID-19 diagnosis (excluding diagnosis date); **4. Hospitalization without ICU**, patients hospitalized with no

record of ICU admission; **5. Hospitalized with ICU but no ventilation**, patients hospitalized with record of ICU admission but no record of ventilator or extracorporeal membrane oxygenation (ECMO) use during ICU stay; **6. Hospitalized with ICU and ventilation**, patients hospitalized with record of ICU admission and ventilator or ECMO use during ICU stay.

Hospitalization was defined as an inpatient or ER overnight visit with an initial COVID-19 diagnosis made during hospitalization and within seven days of admission, or an inpatient or ER overnight visit within 21 days of the initial COVID-19 diagnosis, where the hospital had a record of this diagnosis. Contiguous ER and inpatient visits with a gap of up to one day were considered a single hospitalization. If a patient had multiple eligible hospitalizations, only data from the first hospitalization were considered, as described previously.[17]

Modeling and statistical analysis

LTOs occurring >30–≤180 days after hospital discharge or COVID-19 diagnosis were categorized into one of two time windows (>30–≤90 days or >90–≤180 days) and were further classified as respiratory, cardiovascular, or mental health conditions (**Supplemental Table 1**).[19] LTOs were selected to capture a broad range of potential sequelae, even if there was no strong clinical or pathological rationale for their choice, given the absence of sufficient clinical data regarding established complications associated with COVID-19. Multivariate logistic regression analyses were performed to determine the effect of disease severity (proxied by treatment setting) on the three LTO classifications. Covariates were intended to encompass the main known risk factors for developing severe COVID-19,[20] and included demographic information (i.e., age, gender, race, ethnicity, diagnosis month, insurance type, obesity status) and baseline health conditions (i.e., those included in the Charlson Comorbidity Index [CCI] (**Supplemental Table 2**). CCI was treated as a numeric variable, while all other variables were treated as categorical. Age was binned into <18 years, 18–29, 30–39, 40–49, 50–65, 65–74, 75–84, ≥85 years. Date of diagnosis was also

binned into months in 2020 (pre-April, April, May, June, July; allowing for ≥180 days follow-up until 31 December 2020 at the latest). Patients were excluded from the regression model examining a specific LTO category if they had a diagnosis in that category in the 12 months prior to the index date (for example, if a patient had an asthma diagnosis 12 months prior to the index date, they would be excluded from the model for respiratory LTOs).

All statistical analysis was performed using R 3.6.3.[21] Using the sjstats package, regression was performed using the function 'glm' and the risk ratio (RR) was calculated by converting the odds ratio (OR) using the function 'OR to RR'.[22] Increased risk of diagnosis of a health condition was implied when the RR and both the low and high 95% confidence interval limits (CI) were >1, and decreased risk was implied when the RR and low and high 95% CIs were <1.

Sensitivity analysis

A sensitivity analysis was performed to investigate the potential effect of disease severity (proxied by treatment setting) on risk of a new cancer diagnosis, to serve as a negative control. The same set of covariates was used as per the main analysis, but cancer diagnosis was the only LTO examined. Currently, no evidence exists to suggest that COVID-19 severity increases the risk of a new cancer diagnosis. Thus, an effect here may indicate that the effects from the main analysis may be driven by other differences between patients across treatment settings.

Patient and Public Involvement

No patient involved.

Results

Patient population

In total, 57,748 patients were eligible for the overall COVID-19 cohort. **Table 1** presents descriptive statistics of the patients by sub-cohort. Mean age tended to be higher in patients in hospitalized sub-cohorts (53.2–57.7 years) than in those in non-hospitalized sub-cohorts (41.0–46.8 years). Overall, 53.3% of patients were female. Across all patients, 50.3% were Caucasian, 22.8% were African American, 3.2% were Asian, and the remaining 23.6% were missing information on race. Additionally, 67.5% were of non-Hispanic ethnicity, while data on ethnicity was missing for 11.8% of patients. Overall, 19% of patients were obese and the mean weighted CCI score was 1.20. Information on smoking status was missing for 93.1% of patients (**Table 1**). Full details of demographics and baseline characteristics are provided in **Supplemental Table 3**.

The proportions of patients with incipient respiratory, cardiovascular, and / or mental health conditions that were diagnosed either >30–≤90 days or >90–≤180 days after COVID-19 diagnosis or hospital discharge are provided in **Table 2**. The proportions of patients with new LTOs were generally higher in the sub-cohorts with more severe disease (i.e., the ER sub-cohort and all hospitalized sub-cohorts) compared with the outpatient sub-cohort. In addition, the proportion of patients with respiratory LTOs was higher than the proportions with cardiovascular or mental health LTOs. New respiratory LTOs were diagnosed more frequently during the earlier time window across sub-cohorts, except in the outpatient sub-cohort where the proportion of patients diagnosed was the same in both time windows (both 8.1%; **Table 2**). No clear temporal trends were noted for diagnosis of cardiovascular or mental health LTOs, with similar proportions of patients with new cardiovascular and mental health LTOs observed in the >30–≤90- and >90–≤180-day windows for each sub-cohort (**Table 2**). The proportions of patients with LTOs in more than one category (i.e., 'respiratory and cardiovascular', 'respiratory and mental health', 'mental health and cardiovascular', or 'respiratory, cardiovascular, and mental health') were lower than the proportions of patients

with LTOs in a single category, suggesting that a diagnosis in one category did not necessarily lead to a diagnosis in another.

Regarding individual conditions, the prevalence of newly diagnosed pneumonia, dyspnea, and respiratory failure in the >90–≤180-day window closely followed the pattern of initial COVID-19 severity, with most cases being diagnosed in the 'ICU with ventilation' sub-cohort (**Supplemental Table 4**). Similarly, although encephalopathy, confusion or disorientation, cardiac arrhythmia, and myocardial infarction were less common, the prevalence of these conditions also increased with increasing COVID-19 severity. Full details of conditions that were diagnosed in the >30–≤90- and >90–≤180-day windows following COVID-19 diagnosis or hospital discharge are provided in **Supplemental Table 4**.

Modeling

The most striking potential covariate associated with increased risk of newly diagnosed respiratory conditions at >30–≤90 days and >90–≤180 days post COVID-19 diagnosis or hospital discharge was increasing severity of illness according to increasing hospitalization severity, utilizing the outpatient sub-cohort as the reference group (**Figure 2** and **Supplemental Table 5**). ICU with ventilation was associated with increased risk of a novel respiratory condition diagnosis compared with the outpatient sub-cohort at >30–≤90 days (RR 2.64, 95% CI 2.27, 3.04) and >90–≤180 days post COVID-19 diagnosis or hospital discharge (RR 1.86, 95% CI 1.55, 2.21); in addition, ICU without ventilation was associated with increased risk during the >30–≤90-day time window (RR 1.69, 95% CI 1.39, 2.03), while ER was associated with increased risk at both >30–≤90 days (RR 1.39, 95% CI 1.17, 1.65) and >90–≤180 days (RR 1.33, 95% CI 1.10, 1.58) post COVID-19 diagnosis or hospital discharge. By contrast, patients with an ER visit on the COVID-19 diagnosis date were less likely than those in the outpatient sub-cohort to be diagnosed with a new respiratory condition at >30–≤90 days (RR 0.64, 95% CI 0.56, 0.74) and 90–180 days post COVID-19 diagnosis or hospital discharge (RR 0.56, 95% CI 0.48, 0.65). Additional covariates associated with increased risk of new respiratory conditions were older patient age and

obesity. A COVID-19 diagnosis during or prior to April 2020 exhibited a non-significant trend towards increased risk of new respiratory condition occurrence compared with later diagnosis, which may reflect changes in treatment algorithms over time. Full results are presented in **Supplemental Table 5**.

Increasing hospitalization severity was also found to be associated with increased risk of a new cardiovascular condition occurring post COVID-19 diagnosis or hospital discharge (**Figure 3** and **Supplemental Table 5**). Notably, ICU with ventilation was associated with increased risk of the occurrence of novel cardiovascular conditions compared with the outpatient sub-cohort at >30–≤90 days (RR 3.16, 95% CI 1.83, 5.18) and >90–≤180 days post COVID-19 diagnosis or hospital discharge (RR 2.65, 95% CI 1.49, 4.43), while ICU without ventilation was associated with increased risk during the >90–≤180-day time window (RR 2.41, 95% CI 1.25, 4.23). Similar to the findings regarding respiratory conditions, patients with an ER visit on the COVID-19 diagnosis date were less likely than outpatients to be diagnosed with novel cardiovascular conditions in both the >30–≤90-day (RR 0.45, 95% CI 0.27, 0.71) and >90–≤180-day windows (RR 0.59, 95% CI 0.38, 0.89). Additional covariates associated with an increased risk of new cardiovascular conditions occurring included older patient age and non-Hispanic ethnicity. Full results are presented in **Supplemental Table 5**.

The risk of a new mental health condition occurring post COVID-19 diagnosis or hospital discharge also increased according to increasing hospitalization severity (**Figure 4** and **Supplemental Table 5**). ICU with ventilation was associated with increased risk of a new mental health condition occurring compared with the outpatient sub-cohort at >30–≤90 days (RR 1.89, 95% CI 1.51, 2.35) and >90–≤180 days post COVID-19 diagnosis or hospital discharge (RR 1.52, 95% CI 1.20, 1.91), and ICU without ventilation was similarly associated with increased risk of a new mental health condition diagnosis during the >90–≤180-day window (RR 1.34, 95% CI 1.02, 1.73). Of note, compared with those <18 years, all age groups examined appeared to be at higher risk of the occurrence of new mental health conditions at >30–≤90 days post COVID-19 diagnosis or hospital discharge. In the >90–

≤180-day window, only the 65–74 and 75–84 years age groups were not at higher risk. Additional covariates associated with increased risk of a new mental health condition occurring included obesity, Caucasian race, and non-Hispanic ethnicity. See **Supplemental Table 5** for full results.

Sensitivity analysis

With the exception of older age, COVID-19 severity did not predict a new cancer diagnosis up to 180 days after COVID-19 diagnosis or hospital discharge (**Supplemental Figure 1** and **Supplemental Table 6**), giving confidence in the results of the original analysis.

Discussion

By utilizing EHRs of over 55,000 patients from hospitals and clinics across the US, this study set out to examine the types of new LTOs (i.e., only those that were identified after COVID-19 diagnosis or hospital discharge) associated with long COVID and to identify potential underlying factors that may contribute to their occurrence. Severe disease was found to predict an increased likelihood of a new LTO diagnosis, whereby increasing hospitalization severity was associated with increased risk of new respiratory (e.g., pneumonia), cardiovascular (e.g., myocardial infarction), and mental health conditions (e.g., confusion or disorientation). In severely affected COVID-19 patients, some LTOs were diagnosed between three and six months after hospital discharge, suggesting that the overall COVID-19 burden extends far beyond the acute infection phase. In addition, although patients with severe disease were most at risk of presenting with new LTOs, non-hospitalized patients also experienced a relatively high incidence of LTOs, suggesting that even patients with mild disease are at risk of adverse long-term effects associated with COVID-19.

Although the data show a clear general trend of increased LTOs that correlated with COVID-19 severity, the specificity of this effect to COVID-19 is unclear, as ICU survivors commonly develop a range of new conditions upon discharge collectively referred to as 'post-intensive care syndrome', regardless of their underlying diagnosis.[11] Nonetheless, preventing the development of more severe disease, where possible, may decrease the likelihood of health problems post infection and would be expected to simultaneously increase the probability of survival. Together, these effects would have a cumulative positive impact on both patients and healthcare systems.

Interestingly, the 'ER on diagnosis' sub-cohort exhibited a reduced incidence of LTOs compared with the outpatient sub-cohort. The reasons for this are not clear but are likely due in part to the lower mean age and reduced incidence of comorbidities in this sub-cohort relative to the other sub-cohorts. In addition, it is possible that in the context of the pandemic, when primary care physicians had more limited personal protective equipment

and other resources, these patients were directed to the ER to be tested for COVID-19, despite not having severe enough disease to warrant an ER visit. Finally, depending on the hospital setting and processes in place, asymptomatic patients who attended the ER for non-COVID-19 reasons may have tested positive while there, which may have led to the inclusion of milder COVID-19 cases in this sub-cohort.

Previous studies have examined the link between COVID-19 severity and LTOs. A study of 2,469 hospitalized COVID-19 patients in Wuhan, China showed that more severe disease correlated with increased risk of LTOs up to six months after infection, including fatigue, sleep difficulties, and anxiety or depression.[23] Anxiety or depression was observed in 23% of patients in that study compared with ~10% in our study; this difference is likely because our study was limited to newly diagnosed disorders in both inpatients and outpatients, while the previous study included new or worsening symptoms in hospitalized patients only. A separate, large study of COVID-19 patients that utilized a US EHR database (N=236,379) to examine six-month outcomes (inpatients and outpatients) reported that ~7% of patients had a first anxiety disorder compared with ~17% that had any anxiety disorder, and that increased incidence was correlated with increased disease severity.[13] A further study compared 73,435 non-hospitalized COVID-19 patients who were users of the Veterans Health Administration with 4,990,835 control patients and reported an increased risk of incident sequelae including, but not limited to, respiratory, cardiovascular, and mental health disorders after a median follow-up duration of 126 and 130 days, respectively.[19] Smaller, single-site hospital studies in the United Kingdom have reported similar trends between disease severity and shorter-term outcomes, with breathlessness commonly reported up to 12 weeks post COVID-19.[24 25] In addition, self-reported data in patients with COVID-19 (N=4,182) showed that upper respiratory complaints (e.g., shortness of breath) and cardiac symptoms (e.g., palpitations, tachycardia) were commonly reported in patients with long COVID (symptoms lasting ≥28 days),[26] and data from a separate study utilizing wearable devices provided further evidence of prolonged tachycardia in symptomatic patients with COVID-19.[27] The current study builds on these previous reports and provides additional

evidence of a link between COVID-19 severity and increased risk of developing LTOs, using a large dataset from both hospitalized and non-hospitalized patients. In addition, our study provides a detailed summary of the incidence of a wide range of specific health conditions that occurred up to six months after COVID-19 diagnosis or hospital discharge, providing a useful resource to better understand and characterise the range of conditions that constitute long COVID.

Our study categorized three major classes of LTOs that occur in patients with long COVID: respiratory, cardiovascular, and mental health. This is broadly in keeping with a previous retrospective cohort study in England that followed 48,780 patients hospitalized with COVID-19, who had significantly higher rates of respiratory and cardiovascular disease after a mean follow-up of 140 days.[28] In addition, a retrospective study that used a large administrative all-payor database including 27,589 inpatients and 46,857 outpatients demonstrated that post COVID-19, patients were more likely to experience a range of conditions, including respiratory, nervous, and circulatory system conditions, than outpatient control patients.[29] A greater understanding of the conditions that characterize long COVID is needed to better anticipate the future healthcare burden of COVID-19 and to optimize strategies to minimize long COVID development. In this regard, signals detected in the current study such as lung fibrosis, as well as other factors including pediatric long COVID, vaccination effects, and healthcare utilization, are topics that may warrant future analysis. In particular, a greater understanding of the long-term economic consequences of COVID-19 and the impact of long COVID on patient quality of life is needed.

A major limitation of this analysis is that treatment setting is used as a proxy for COVID-19 severity; therefore, it is difficult to tease out the effect of treatment setting procedures (e.g., invasive ventilation) from the underlying COVID-19 severity. Furthermore, our analysis did not distinguish short-term from chronic health conditions. Additional limitations include missing information on smoking status, the restriction of follow-up to only six months, the lack of a COVID-19-negative control group, the possibility of missing data

(e.g., patients may have sought care for an LTO not captured in the Optum® de-identified COVID-19 EHR dataset), the lack of information on COVID-19 treatments received, and the lack of laboratory values or other biomarkers to better characterize disease. Finally, capture of health conditions relies on International Classification of Disease-10 (ICD-10) codes, whereas some conditions of interest (e.g., anosmia, ageusia, and brain fog) lack specific ICD-10 codes and other conditions are known to be under-captured. The B97.29 diagnosis code includes other coronaviruses in addition to SARS-CoV-2 and may therefore be a potential limitation of our study; however, the majority of our COVID-19 cohort (>85%) was diagnosed from April to July using the official U07.1 diagnosis code that is specific to COVID-19, meaning it is unlikely that a substantial number of infections, if any, were from other coronaviruses.

Conclusions

Although LTOs were reported in patients across all sub-cohorts, increased risk of new respiratory, cardiovascular, and mental health conditions was observed with increasing COVID-19 severity. Strikingly, the risk of new conditions being diagnosed remained high up to six months post COVID-19 diagnosis or hospital discharge, suggesting that the burden of COVID-19 extends far beyond the acute infection phase. Future research is warranted to understand specific factors that lead to the occurrence of new LTOs in patients with COVID-19, and to distinguish between the relative effect of COVID-19 severity versus any general effects that may occur after acute critical illness.

Acknowledgements

We thank Shemra Rizzo for valuable contributions. Third-party medical writing assistance, under the direction of the authors, was provided by John Bett, PhD, of Ashfield MedComms, an Ashfield Health company, and funded by F. Hoffmann-La Roche Ltd.

Contributors

All authors were involved in drafting and revising the manuscript, approved the final version, and agree to being accountable for all aspects of the work. Nick Jovanoski contributed to the conception of the research question, study design, analysis, and data interpretation. Xin Chen contributed to study design, analysis and data interpretation. Ursula Becker contributed to the conception of the research question, study design, analysis, and data interpretation. Kelly Zalocusky contributed to the conception of the research question, design of the analysis, selection of outcomes and data interpretation. Devika Chawla contributed to the conception of the research question, design of the analysis, and selection of outcomes. Larry Tsai contributed to study design and data interpretation. Michelle Borm contributed to data interpretation. Margaret Neighbors contributed to selection and categorization of key complications for study design. Vincent Yau contributed to the study design, acquisition, analysis and data interpretation.

Funding statement

This work was supported by F. Hoffmann-La Roche Ltd.

Competing interests statement

Nick Jovanoski and Ursula Becker are employees of F. Hoffmann-La Roche Ltd. Michelle Borm is an employee of Roche Nederland BV. Ursula Becker and Michelle Borm hold shares in F. Hoffmann-La Roche Ltd. Xin Chen, Devika Chawla, Larry Tsai, Margaret Neighbors, and Vincent Yau are employees of Genentech, Inc. and hold shares in F. Hoffmann-La

Roche Ltd. Kelly Zalocusky is a former employee of Genentech, Inc. and holds shares in F. Hoffmann-La Roche Ltd.

Patient consent

None required

Ethics approval

The use of the Optum® de-identified COVID-19 EHR dataset was reviewed by the New England Institutional Review Board (IRB) and was determined to be exempt from broad IRB approval, as this study did not involve human subject research.

Data availability statement

Data may be obtained from a third party and are not publicly available. Data were licensed from Optum® and interested researchers may contact Optum® for data access requests. All interested researchers can access the data in the same manner as the authors. The authors had no special access privileges.

References

1. Centers for Disease Control and Prevention. Estimated Disease Burden of COVID-19 2021 [Available from: https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/burden.html accessed May 2021.
2. Carfi A, Bernabei R, Landi F, et al. Persistent Symptoms in Patients After Acute COVID-19. *JAMA* 2020;324(6):603-05. doi: 10.1001/jama.2020.12603 [published Online First: 2020/07/10]
3. Tenforde MW, Kim SS, Lindsell CJ, et al. Symptom Duration and Risk Factors for Delayed Return to Usual Health Among Outpatients with COVID-19 in a Multistate Health Care Systems Network - United States, March-June 2020. *MMWR Morb Mortal Wkly Rep* 2020;69(30):993-98. doi: 10.15585/mmwr.mm6930e1 [published Online First: 2020/07/31]
4. Nalbandian A, Sehgal K, Gupta A, et al. Post-acute COVID-19 syndrome. *Nat Med* 2021;27(4):601-15. doi: 10.1038/s41591-021-01283-z [published Online First: 2021/03/24]
5. Centers for Disease Control and Prevention. Post-COVID Conditions 2021 [Available from: https://www.cdc.gov/coronavirus/2019-ncov/long-term-effects.html accessed May 2021.
6. Augustin M, Schommers P, Stecher M, et al. Post-COVID syndrome in non-hospitalised patients with COVID-19: a longitudinal prospective cohort study. *Lancet Reg Health Eur* 2021;6:100122. doi: 10.1016/j.lanepe.2021.100122 [published Online First: 2021/05/25]
7. Mullard A. Long COVID's long R&D agenda. *Nat Rev Drug Discov* 2021;20(5):329-31. doi: 10.1038/d41573-021-00069-9 [published Online First: 2021/04/22]
8. National Institutes of Health. NIH launches new initiative to study “Long COVID” Bethesda, Maryland: National Institutes of Health; 2021 [Available from: https://www.nih.gov/about-nih/who-we-are/nih-director/statements/nih-launches-new-initiative-study-long-covid accessed July 2021.

9. Mayor N, Tsang R, Joy M, et al. Long covid: coding is caring. *BMJ* 2021;373:n1262. doi: 10.1136/bmj.n1262 [published Online First: 2021/05/21]

10. Marx V. Scientists set out to connect the dots on long COVID. *Nat Methods* 2021;18(5):449-53. doi: 10.1038/s41592-021-01145-z [published Online First: 2021/04/30]

11. Brown SM, Bose S, Banner-Goodspeed V, et al. Approaches to Addressing Post-Intensive Care Syndrome among Intensive Care Unit Survivors. A Narrative Review. *Ann Am Thorac Soc* 2019;16(8):947-56. doi: 10.1513/AnnalsATS.201812-913FR [published Online First: 2019/06/05]

12. Myall KJ, Mukherjee B, Castanheira AM, et al. Persistent Post-COVID-19 Interstitial Lung Disease. An Observational Study of Corticosteroid Treatment. *Ann Am Thorac Soc* 2021;18(5):799-806. doi: 10.1513/AnnalsATS.202008-1002OC [published Online First: 2021/01/13]

13. Taquet M, Geddes JR, Husain M, et al. 6-month neurological and psychiatric outcomes in 236 379 survivors of COVID-19: a retrospective cohort study using electronic health records. *Lancet Psychiatry* 2021;8(5):416-27. doi: 10.1016/S2215-0366(21)00084-5 [published Online First: 2021/04/10]

14. Puntmann VO, Carerj ML, Wieters I, et al. Outcomes of Cardiovascular Magnetic Resonance Imaging in Patients Recently Recovered From Coronavirus Disease 2019 (COVID-19). *JAMA Cardiol* 2020;5(11):1265-73. doi: 10.1001/jamacardio.2020.3557 [published Online First: 2020/07/31]

15. University of Birmingham. Why are some people more susceptible to the long-term health effects of COVID-19? 2021 [Available from: https://www.birmingham.ac.uk/research/quest/21st-century-healthcare/long-covid.aspx accessed May 2021.

16. Mount Sinai Hospital. Center for Post-COVID Care 2021 [Available from: https://www.mountsinai.org/about/covid19/center-post-covid-care accessed May 2021.

17. Chawla D, Rizzo S, Zalocusky K, et al. Descriptive epidemiology of 16,780 hospitalized COVID-19 patients in the United States. *medRxiv* 2021 doi: https://doi.org/10.1101/2020.07.17.20156265

18. Incerti D, Rizzo S, Li X, et al. Prognostic model to identify and quantify risk factors for mortality among hospitalised patients with COVID-19 in the USA. *BMJ open* 2021;11(4):e047121.

19. Al-Aly Z, Xie Y, Bowe B. High-dimensional characterization of post-acute sequelae of COVID-19. *Nature* 2021 doi: 10.1038/s41586-021-03553-9

20. Centers for Disease Control and Prevention. People with Certain Medical Conditions 2021 [Available from: https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/people-with-medical-conditions.html accessed July 2021.

21. R Core Team. R: A language and environment for statistical computing Vienna, Austria2020 [Available from: https://www.r-project.org/.

22. Lüdecke D. Statistical Functions for Regression Models (Version 0.18.1) 2021 [Available from: https://CRAN.R-project.org/package=sjstats.

23. Huang C, Huang L, Wang Y, et al. 6-month consequences of COVID-19 in patients discharged from hospital: a cohort study. *Lancet* 2021;397(10270):220-32. doi: 10.1016/S0140-6736(20)32656-8 [published Online First: 2021/01/12]

24. Halpin SJ, McIvor C, Whyatt G, et al. Postdischarge symptoms and rehabilitation needs in survivors of COVID-19 infection: A cross-sectional evaluation. *J Med Virol* 2021;93(2):1013-22. doi: 10.1002/jmv.26368 [published Online First: 2020/07/31]

25. Arnold DT, Hamilton FW, Milne A, et al. Patient outcomes after hospitalisation with COVID-19 and implications for follow-up: results from a prospective UK cohort. *Thorax* 2020 doi: 10.1136/thoraxjnl-2020-216086 [published Online First: 2020/12/05]

26. Sudre CH, Murray B, Varsavsky T, et al. Attributes and predictors of long COVID. *Nature Medicine* 2021;27(4):626-31. doi: 10.1038/s41591-021-01292-y

27. Radin JM, Quer G, Ramos E, et al. Assessment of Prolonged Physiological and Behavioral Changes Associated With COVID-19 Infection. *JAMA Netw Open*

2021;4(7):e2115959. doi: 10.1001/jamanetworkopen.2021.15959 [published Online First: 2021/07/08]

28. Ayoubkhani D, Khunti K, Nafilyan V, et al. Post-covid syndrome in individuals admitted to hospital with covid-19: retrospective cohort study. *BMJ* 2021;372:n693. doi: 10.1136/bmj.n693 [published Online First: 2021/04/02]
29. Chevinsky JR, Tao G, Lavery AM, et al. Late conditions diagnosed 1-4 months following an initial COVID-19 encounter: a matched cohort study using inpatient and outpatient administrative data - United States, March 1-June 30, 2020. *Clin Infect Dis* 2021 doi: 10.1093/cid/ciab338 [published Online First: 2021/04/29]

Tables and Figures

Table 1 Baseline characteristics of COVID-19 patients overall and by sub-cohort

		Sub-cohort					
	All patients (N=57,748)	**1. Outpatient (n=22,788)**	**2. ER on diagnosis (n=11,633)**	**3. ER (n=2,877)**	**4. Hospitalization without ICU (n=16,653)**	**5. ICU without ventilation (n=1,837)**	**6. ICU with ventilation (n=1,960)**
Mean age (SD), years	47.93 (18.76)	46.78 (18.90)	40.95 (16.89)	44.00 (16.61)	53.17 (18.50)	55.92 (17.32)	57.70 (14.78)
Age group, n (%)							
<18 years	2,184 (3.8)	1,033 (4.5)	641 (5.5)	78 (2.7)	366 (2.2)	46 (2.5)	20 (1.0)
18–29 years	8,509 (14.7)	3,693 (16.2)	2,543 (21.9)	538 (18.7)	1,574 (9.5)	89 (4.8)	72 (3.7)
30–39 years	8,972 (15.5)	3,541 (15.5)	2,496 (21.5)	579 (20.1)	2,046 (12.3)	175 (9.5)	135 (6.9)
40–49 years	9,362 (16.2)	3,664 (16.1)	2,205 (19.0)	555 (19.3)	2,442 (14.7)	253 (13.8)	243 (12.4)
50–64 years	16,103 (27.9)	6,231 (27.3)	2,706 (23.3)	780 (27.1)	4,937 (29.6)	626 (34.1)	823 (42.0)
65–74 years	7,065 (12.2)	2,658 (11.7)	689 (5.9)	230 (8.0)	2,683 (16.1)	363 (19.8)	442 (22.6)
75–84 years	3,620 (6.3)	1,279 (5.6)	239 (2.1)	76 (2.6)	1,647 (9.9)	195 (10.6)	184 (9.4)
≥85 years	891 (1.5)	303 (1.3)	54 (0.5)	22 (0.8)	440 (2.6)	44 (2.4)	28 (1.4)
Missing	1,042 (1.8)	386 (1.7)	60 (0.5)	19 (0.7)	518 (3.1)	46 (2.5)	13 (0.7)
Sex, n (%)							

Female	30,782 (53.3)	12,856 (56.4)	6,115 (52.6)	1,721 (59.8)	8,487 (51.0)	829 (45.1)	774 (39.5)
Male	26,939 (46.6)	9,920 (43.5)	5,515 (47.4)	1,152 (40.0)	8,160 (49.0)	1,008 (54.9)	1,184 (60.4)
Missing	27 (<0.1)	12 (<0.1)	3 (<0.1)	4 (<0.1)	6 (<0.1)	0 (0.0)	2 (<0.1)
Race, n (%)							
African American	13,183 (22.8)	3,473 (15.2)	3,178 (27.3)	790 (27.5)	4,675 (28.1)	551 (30.0)	516 (26.3)
Asian	1,848 (3.2)	639 (2.8)	438 (3.8)	100 (3.5)	555 (3.3)	41 (2.2)	75 (3.8)
Caucasian	29,074 (50.3)	13,746 (60.3)	4,653 (40.0)	1,337 (46.5)	7,538 (45.3)	849 (46.2)	951 (48.5)
Missing	13,643 (23.6)	4,930 (21.6)	3,364 (28.9)	650 (22.6)	3,885 (23.3)	396 (21.6)	418 (21.3)
Ethnicity, n (%)							
Hispanic	11,932 (20.7)	3,942 (17.3)	3,378 (29.0)	646 (22.5)	3,298 (19.8)	332 (18.1)	336 (17.1)
Non-Hispanic	38,988 (67.5)	15,485 (68.0)	7,121 (61.2)	1,987 (69.1)	11,648 (69.9)	1,294 (70.4)	1,453 (74.1)
Missing	6,828 (11.8)	3,361 (14.7)	1,134 (9.7)	244 (8.5)	1,707 (10.3)	211 (11.5)	171 (8.7)
Smoking status, n (%)							
Current smoker	413 (0.7)	193 (0.8)	128 (1.1)	13 (0.5)	61 (0.4)	13 (0.7)	5 (0.3)
Previously smoked	740 (1.3)	417 (1.8)	111 (1.0)	27 (0.9)	145 (0.9)	25 (1.4)	15 (0.8)
Never smoked	2,831 (4.9)	1,468 (6.4)	750 (6.4)	121 (4.2)	404 (2.4)	50 (2.7)	38 (1.9)
Missing	53,764 (93.1)	20,710 (90.9)	10,644 (91.5)	2,716 (94.4)	16,043 (96.3)	1,749 (95.2)	1,902 (97.0)

Obese, n (%)*	10,952 (19.0)	4,905 (21.5)	1,366 (11.7)	580 (20.2)	3,246 (19.5)	406 (22.1)	449 (22.9)
Insurance, n (%)							
Commercial	29,145 (50.5)	13,134 (57.6)	5,672 (48.8)	1,482 (51.5)	7,243 (43.5)	758 (41.3)	856 (43.7)
Medicaid	8,652 (15.0)	2,341 (10.3)	2,223 (19.1)	542 (18.8)	2,891 (17.4)	312 (17.0)	343 (17.5)
Medicare	8,774 (15.2)	3,173 (13.9)	788 (6.8)	245 (8.5)	3,674 (22.1)	435 (23.7)	459 (23.4)
Other payor type	4,004 (6.9)	1,282 (5.6)	1,071 (9.2)	211 (7.3)	1,188 (7.1)	129 (7.0)	123 (6.3)
Uninsured	4,833 (8.4)	1,542 (6.8)	1,731 (14.9)	281 (9.8)	1,069 (6.4)	111 (6.0)	99 (5.1)
Missing	2,340 (4.1)	1,316 (5.8)	148 (1.3)	116 (4.0)	588 (3.5)	92 (5.0)	80 (4.1)
Month of COVID-19 diagnosis, n (%)							
Feb 2020	115 (0.2)	61 (0.3)	3 (<0.1)	4 (0.1)	34 (0.2)	5 (0.3)	8 (0.4)
Mar 2020	8,197 (14.2)	1,527 (6.7)	1,893 (16.3)	527 (18.3)	3,288 (19.7)	306 (16.7)	656 (33.5)
Apr 2020	18,591 (32.2)	6,018 (26.4)	3,480 (29.9)	926 (32.2)	6,684 (40.1)	676 (36.8)	807 (41.2)
May 2020	14,188 (24.6)	6,154 (27.0)	2,703 (23.2)	729 (25.3)	3,755 (22.5)	477 (26.0)	370 (18.9)
Jun 2020	14,832 (25.7)	7,846 (34.4)	3,073 (26.4)	597 (20.8)	2,826 (17.0)	371 (20.2)	119 (6.1)
Jul 2020	1,825 (3.2)	1,182 (5.2)	481 (4.1)	94 (3.3)	66 (0.4)	2 (0.1)	0 (0.0)
Mean weighted CCI (SD)	1.20 (2.06)	1.09 (2.03)	0.56 (1.32)	0.88 (1.76)	1.64 (2.31)	2.15 (2.46)	2.13 (2.39)

CCI, Charlson Comorbidity Index; COVID-19, coronavirus disease 2019; ER, emergency room; ICU, intensive care unit; SD, standard deviation

*No patient data were missing for obesity

Table 2 Long-term outcomes that were diagnosed >30–≤90 days or >90–≤180 days post COVID-19 by sub-cohort

Condition, n **(%)**	**1. Outpatient (N=22,788)**		**2. ER on diagnosis (N=11,633)**		**3. ER (N=2,877)**		**4. Hospitalization without ICU (N=16,653)**		**5. ICU without ventilation (N=1,837)**		**6. ICU with ventilation (N=1,960)**	
	>30–≤90 days	**>90–≤180 days**	**>30–≤90 days**	**>90–≤180 days**	**>30–≤90 days**	**>90–≤180 days**	**>30–≤90 days**	**>90–≤180 days**	**>30–≤90 days**	**>90–≤180 days**	**>30–≤90 days**	**>90–≤180 days**
Respiratory	1,846 (**8.1**)	1,846 (**8.1**)	520 (**4.5**)	484 (**4.2**)	322 (**11.2**)	311 (**10.8**)	2,124 (**12.8**)	1,599 (**9.6**)	296 (**16.1**)	222 (**12.1**)	547 (**27.9**)	381 (**19.4**)
CV	439 (**1.9**)	475 (**2.1**)	81 (**0.7**)	92 (**0.8**)	54 (**1.9**)	57 (**2.0**)	585 (**3.5**)	618 (**3.7**)	102 (**5.6**)	99 (**5.4**)	128 (**6.5**)	132 (**6.7**)
Mental health	940 (**4.1**)	1,081 (**4.7**)	231 (**2.0**)	293 (**2.5**)	144 (**5.0**)	154 (**5.4**)	819 (**4.9**)	866 (**5.2**)	119 (**6.5**)	118 (**6.4**)	167 (**8.5**)	144 (**7.3**)
Cancer	137 (**0.6**)	151 (**0.7**)	24 (**0.2**)	30 (**0.3**)	18 (**0.6**)	14 (**0.5**)	95 (**0.6**)	106 (**0.6**)	9 (**0.5**)	12 (**0.7**)	22 (**1.1**)	14 (**0.7**)
Respiratory and CV	131 (**0.6**)	124 (**0.5**)	16 (**0.1**)	30 (**0.3**)	17 (**0.6**)	19 (**0.7**)	201 (**1.2**)	181 (**1.1**)	30 (**1.6**)	26 (**1.4**)	63 (**3.2**)	47 (**2.4**)
Respiratory and mental health	205 (**0.9**)	216 (**0.9**)	65 (**0.6**)	63 (**0.5**)	48 (**1.7**)	33 (**1.1**)	262 (**1.6**)	245 (**1.5**)	45 (**2.4**)	34 (**1.9**)	83 (**4.2**)	50 (**2.6**)
Mental health and CV	53 (**0.2**)	49 (**0.2**)	8 (**0.1**)	9 (**0.1**)	4 (**0.1**)	2 (**0.1**)	57 (**0.3**)	64 (**0.4**)	10 (**0.5**)	11 (**0.6**)	7 (**0.4**)	11 (**0.6**)
Respiratory, CV, and mental health	57 (**0.3**)	52 (**0.2**)	9 (**0.1**)	10 (**0.1**)	6 (**0.2**)	9 (**0.3**)	98 (**0.6**)	84 (**0.5**)	17 (**0.9**)	15 (**0.8**)	30 (**1.5**)	31 (**1.6**)
No new conditions* (respiratory, CV, or mental health)	20,066 (**88.1**)	19,879 (**87.2**)	10,908 (**93.8**)	10,886 (**93.6**)	2,438 (**84.7**)	2,427 (**84.4**)	13,841 (**83.1**)	14,228 (**85.4**)	1,439 (**78.3**)	1,499 (**81.6**)	1,331 (**67.9**)	1,473 (**75.2**)

COVID-19, coronavirus disease 2019; CV, cardiovascular; ER, emergency room; ICU, intensive care unit

*Only conditions that appeared >30–≤180 days after COVID-19 diagnosis or hospital discharge are included; pre-existing conditions are excluded

Figure 1

Title: Overall study design.

Abbreviations: COVID-19, coronavirus disease 2019

Figure 2

Title: Relative risk of new respiratory conditions occurring from >30–≤180 days after COVID-19 diagnosis or hospital discharge.

Abbreviations: CCI, Charlson Comorbidity Index; COVID-19, coronavirus disease 2019; ER, emergency room; ICU, intensive care unit

Legend: Relative risk of new respiratory conditions occurring at (A) >30–≤90 days and (B) >90–≤180 days after COVID-19 diagnosis or hospital discharge. Graphs represent relative risk and 95% confidence intervals. Reference groups are: <18 years (age), female (sex), African American (race), Hispanic (ethnicity), non-obese (obesity), diagnosis in February and March 2020 (diagnosis month); commercial (insurance), cohort 1: outpatient (cohort).

Figure 3

Title: Relative risk of new cardiovascular conditions occurring from >30–≤180 days after COVID-19 diagnosis or hospital discharge.

Abbreviations: CCI, Charlson Comorbidity Index; COVID-19, coronavirus disease 2019; ER, emergency room; ICU, intensive care unit

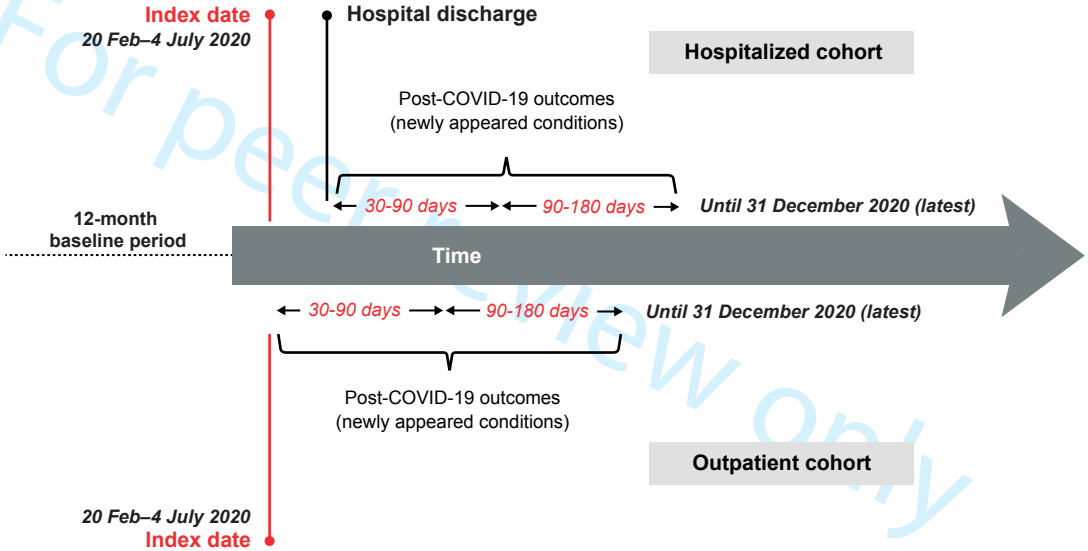
Legend: Relative risk of new cardiovascular conditions occurring at (A) >30–≤90 days and (B) >90–≤180 days after COVID-19 diagnosis or hospital discharge. Graphs represent relative risk and 95% confidence intervals. Reference groups are: <18 years (age), female (sex), African American (race), Hispanic (ethnicity), non-obese (obesity), diagnosis in February and March 2020 (diagnosis month); commercial (insurance), cohort 1: outpatient (cohort). Relative risk in the >30–≤90 days time window was not calculated as no new diagnoses were made in the reference group (<18 years) during this time.

Figure 4

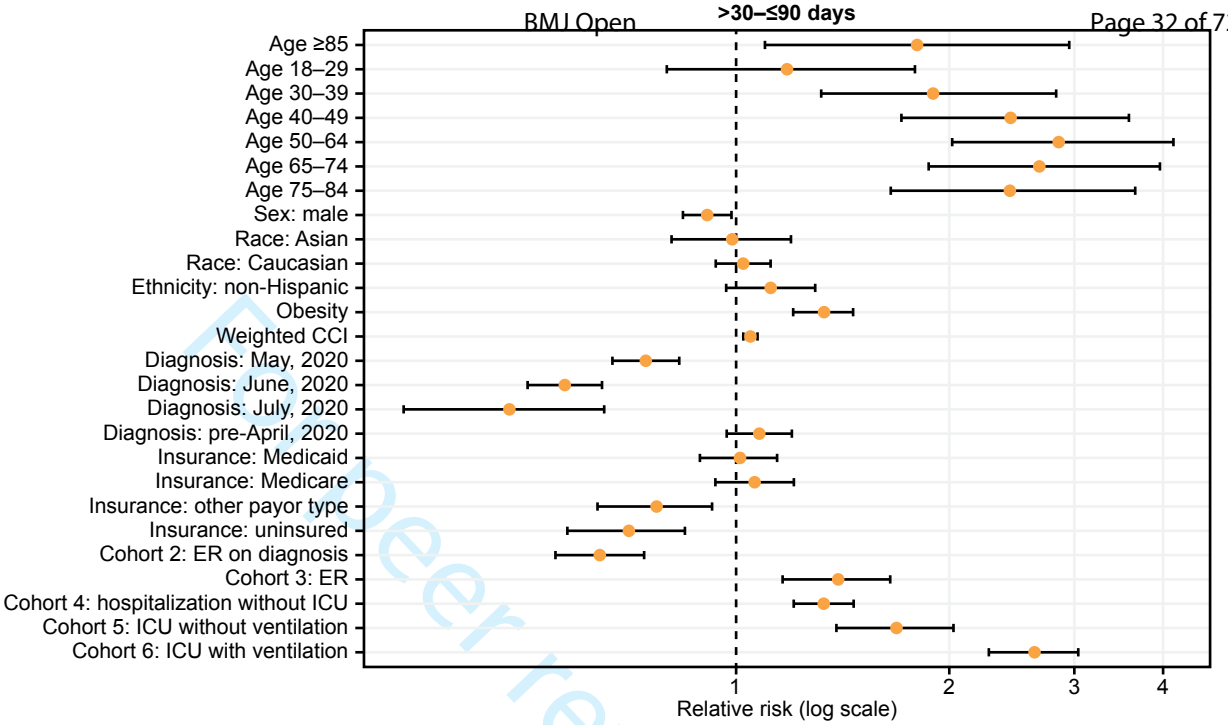
Title: Relative risk of new mental health conditions occurring from >30–≤180 days after COVID-19 diagnosis or hospital discharge.

Abbreviations: CCI, Charlson Comorbidity Index; COVID-19, coronavirus disease 2019; ER, emergency room; ICU, intensive care unit

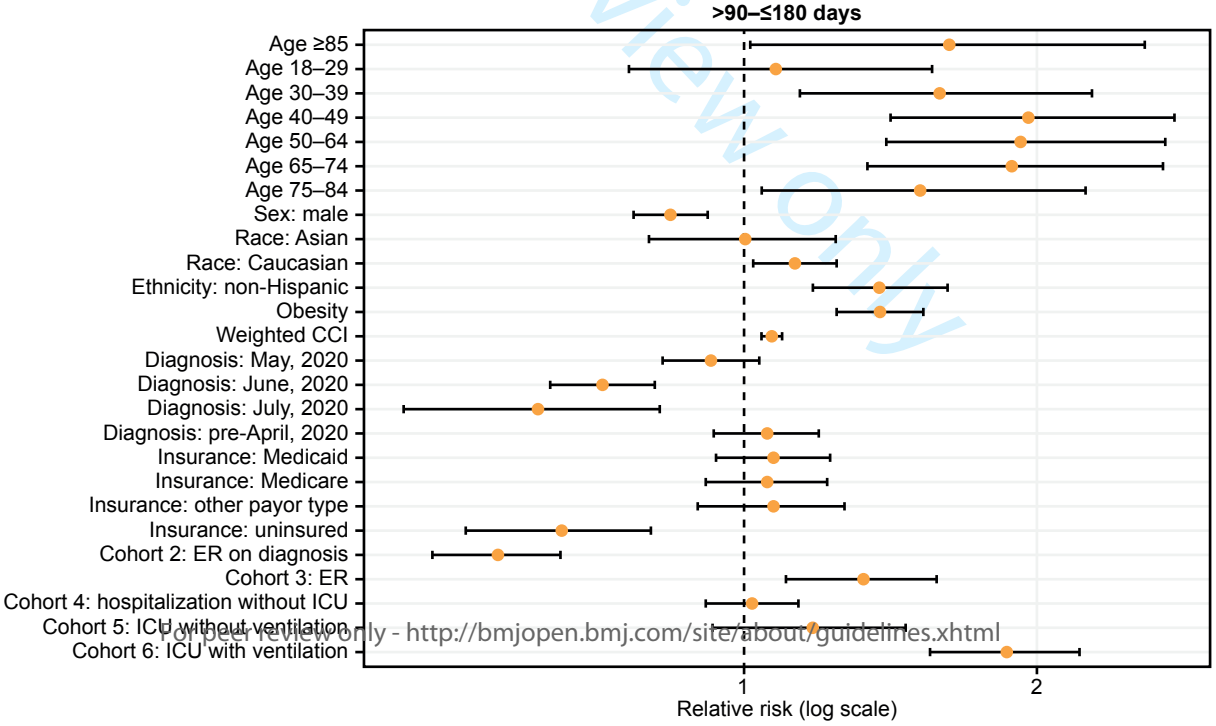
Legend: Relative risk of new mental health conditions occurring at (A) >30–≤90 days and (B) >90–≤180 days after COVID-19 diagnosis or hospital discharge. Graphs represent relative risk and 95% confidence intervals. Reference groups are: <18 years (age), female (sex), African American (race), Hispanic (ethnicity), non-obese (obesity), diagnosis in February and March 2020 (diagnosis month); commercial (insurance), cohort 1: outpatient (cohort).

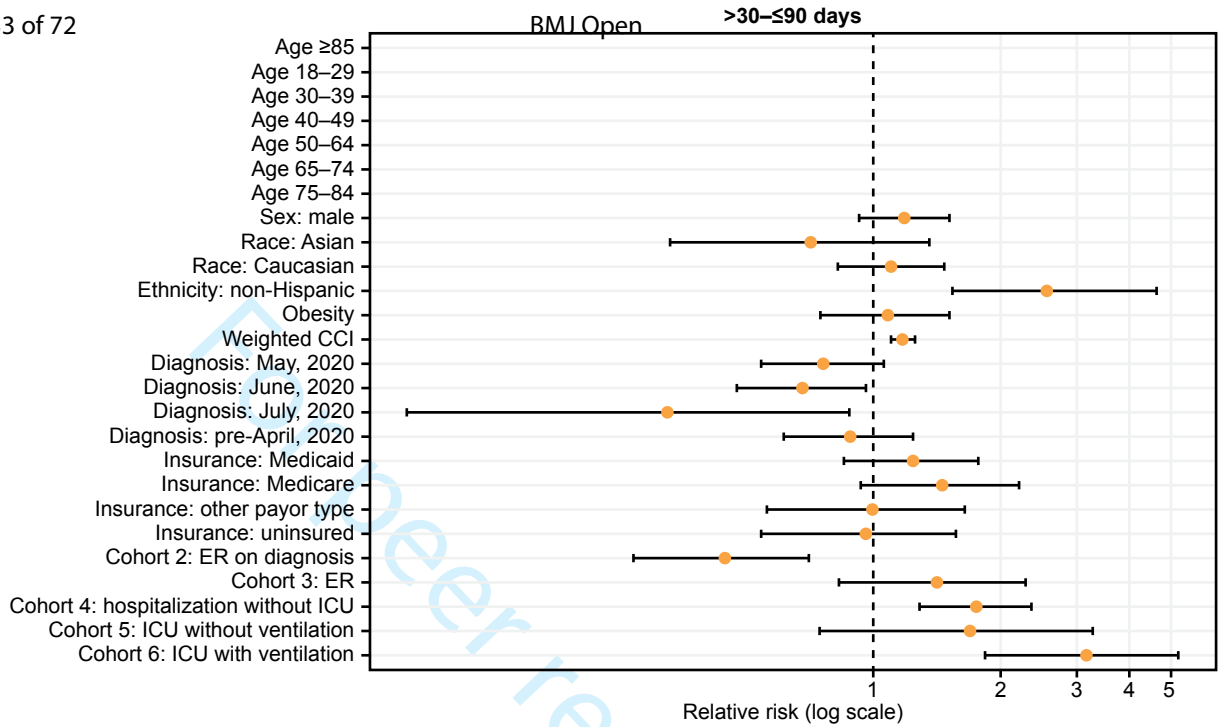


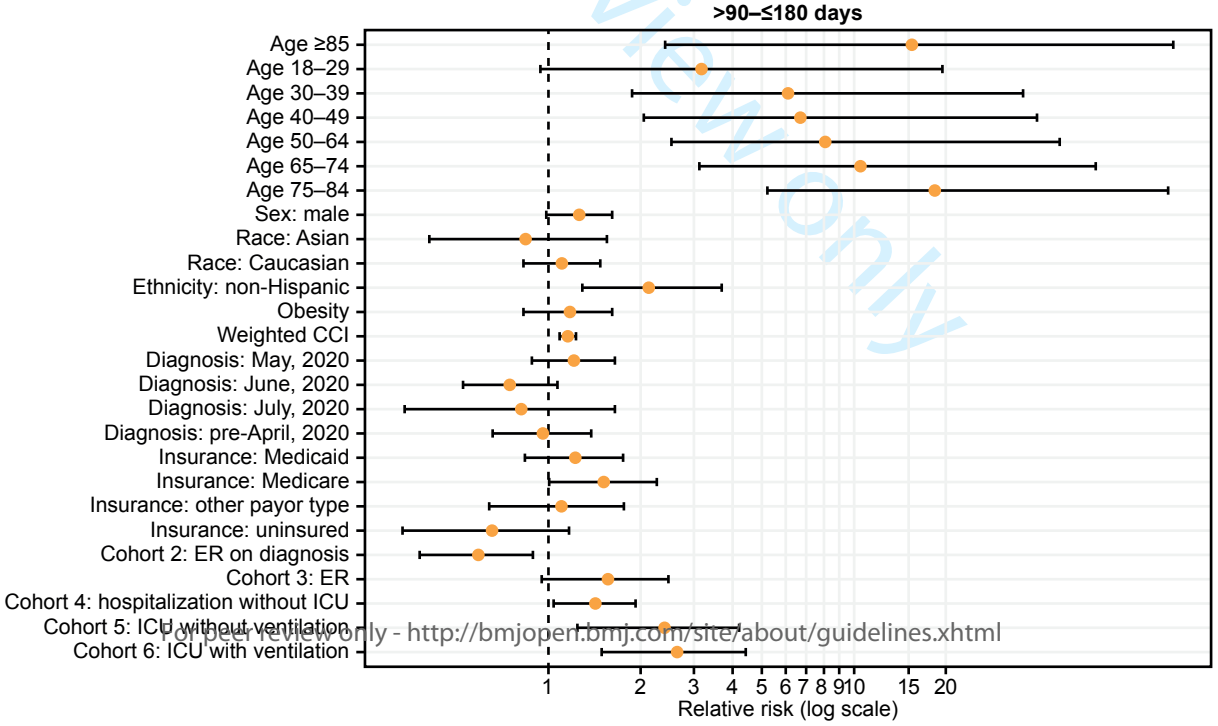
A

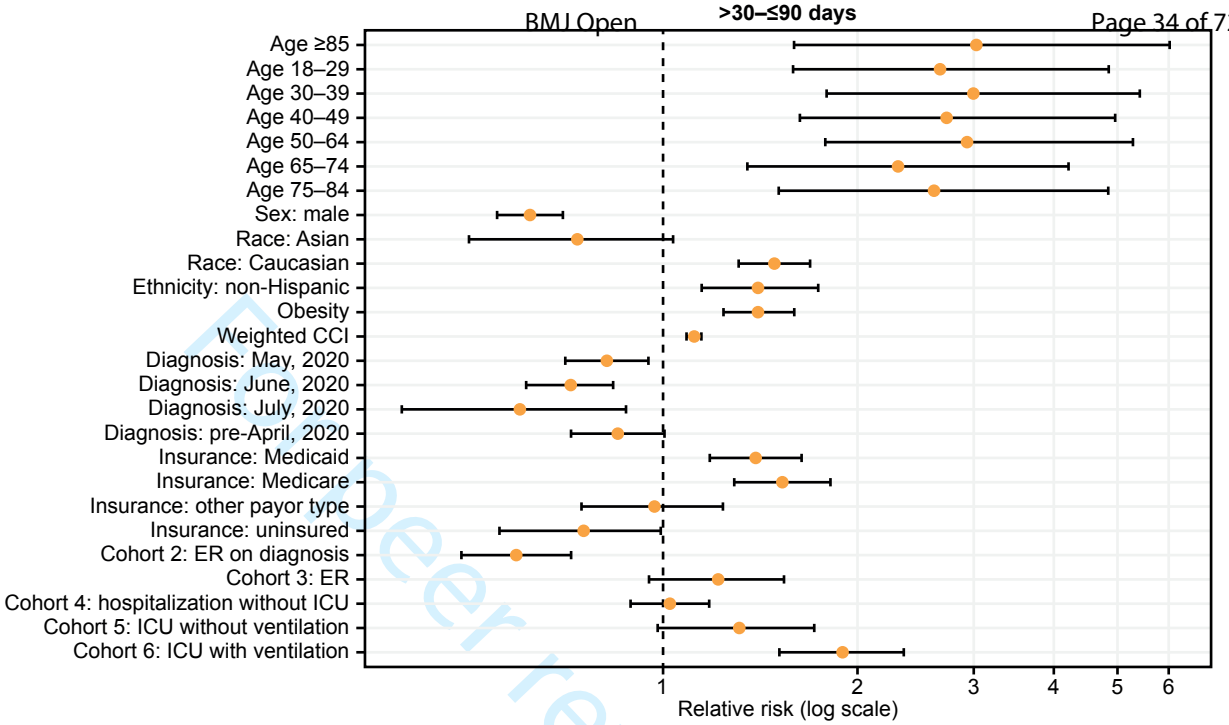


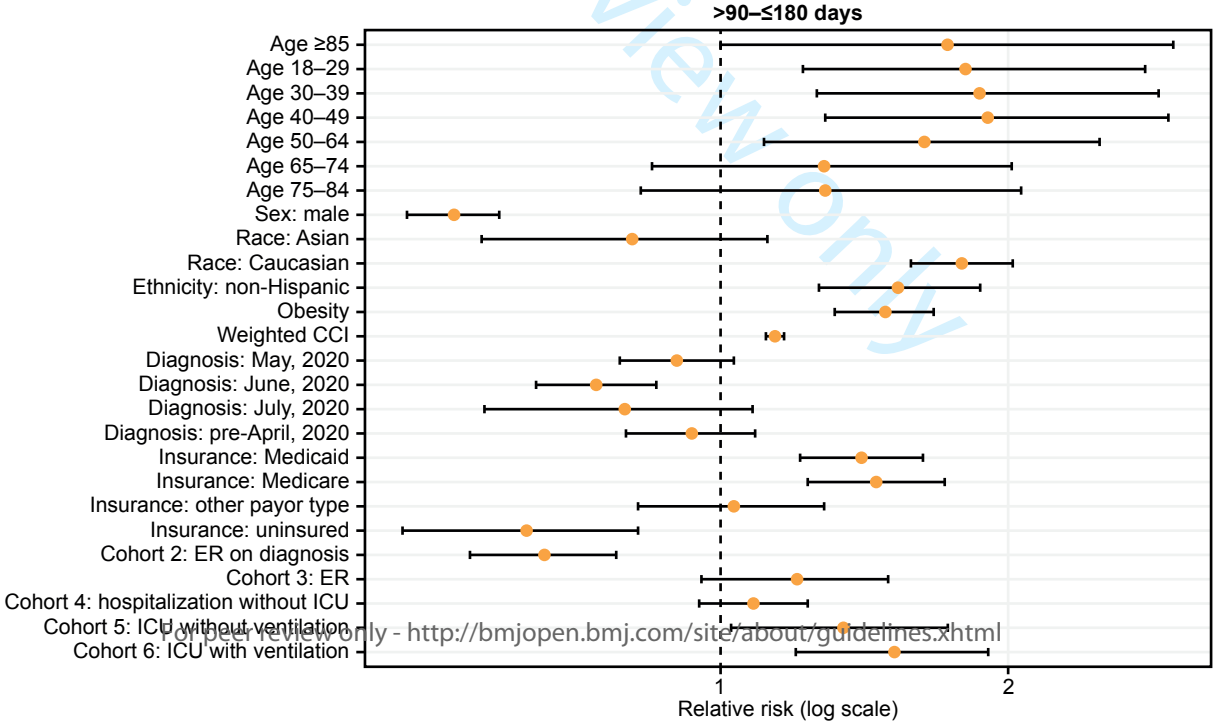
B







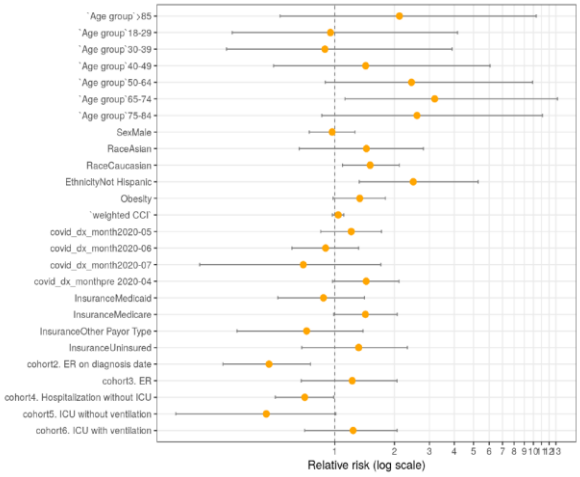




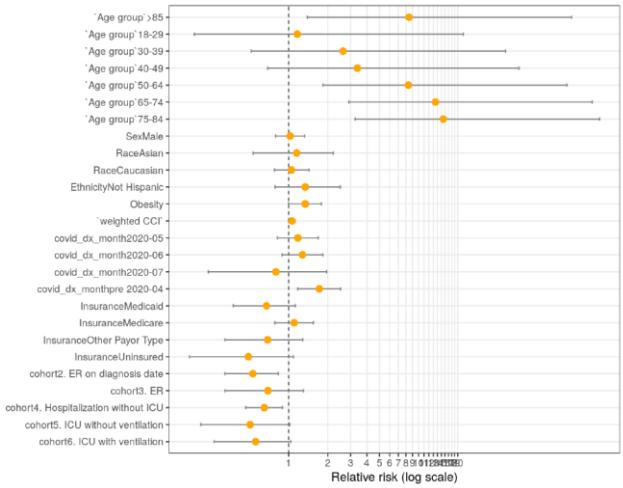
Supplemental materials

Supplemental Figure 1 Relative risk of a new cancer diagnosis from >30–≤180 days after COVID-19 diagnosis or hospital discharge.

A



B



CCI, Charlson Comorbidity Index; COVID-19, coronavirus disease 2019; ER, emergency room; ICU, intensive care unit

Relative risk of a new cancer diagnosis at (A) >30–≤90 days and (B) >90–≤180 days after COVID-19 diagnosis or hospital discharge. Graphs represent relative risk and 95% confidence intervals. Reference groups are: <18 years (age), female (sex), African American (race), Hispanic (ethnicity), non-obese (obesity), diagnosis in February and March 2020 (diagnosis month); commercial (insurance), cohort 1: outpatient (cohort).

Supplemental Table 1 List of the long-term outcomes studied and their classification

Long-term outcome
Respiratory
Asthma
Bronchiectasis
Bronchitis
COPD
Dyspnea
Emphysema
Influenza
Interstitial lung disease (fibrosis)
Pneumonia
Respiratory failure
Cardiovascular
Cardiac arrhythmia
Myocardial infarction
Pulmonary embolism
Pulmonary hypertension
Stroke
Mental health
Anxiety
Confusion or disorientation
Dementia
Depression
Encephalopathy
Memory loss

COPD, chronic obstructive pulmonary disease

Supplemental Table 2 List of comorbidities included in the Charlson Comorbidity Index

Comorbidity	
AIDS	Metastatic solid tumor
Cancer	Mild liver disease
Cerebrovascular disease	Moderate or severe liver disease
Chronic pulmonary disease	Moderate or severe renal disease
Congestive heart failure	Myocardial infarction
Dementia	Peptic ulcer disease
Diabetes with complication	Peripheral vascular disease
Diabetes without complication	Rheumatics
Hemiplegia	

AIDS, acquired immunodeficiency syndrome

Supplemental Table 3 Full baseline characteristics, COVID-19-related outcomes, symptoms, and tests

	All patients (N=57,748)	1. Outpatient (n=22,788)	2. ER on diagnosis (n=11,633)	3. ER (n=2,877)	4. Hospitalization without ICU (N=16,653)	5. ICU without ventilation (N=1,837)	6. ICU with ventilation (N=1,960)
Mean age (SD), years	47.93 (18.76)	46.78 (18.90)	40.95 (16.89)	44.00 (16.61)	53.17 (18.50)	55.92 (17.32)	57.70 (14.78)
Age group, n (%)							
<18 years	2,184 (3.8)	1,033 (4.5)	641 (5.5)	78 (2.7)	366 (2.2)	46 (2.5)	20 (1.0)
18–29 years	8,509 (14.7)	3,693 (16.2)	2,543 (21.9)	538 (18.7)	1574 (9.5)	89 (4.8)	72 (3.7)
30–39 years	8,972 (15.5)	3,541 (15.5)	2,496 (21.5)	579 (20.1)	2,046 (12.3)	175 (9.5)	135 (6.9)
40–49 years	9,362 (16.2)	3,664 (16.1)	2,205 (19.0)	555 (19.3)	2,442 (14.7)	253 (13.8)	243 (12.4)
50–64 years	16,103 (27.9)	6,231 (27.3)	2,706 (23.3)	780 (27.1)	4,937 (29.6)	626 (34.1)	823 (42.0)
65–74 years	7,065 (12.2)	2,658 (11.7)	689 (5.9)	230 (8.0)	2,683 (16.1)	363 (19.8)	442 (22.6)
75–84 years	3,620 (6.3)	1,279 (5.6)	239 (2.1)	76 (2.6)	1,647 (9.9)	195 (10.6)	184 (9.4)
≥85 years	891 (1.5)	303 (1.3)	54 (0.5)	22 (0.8)	440 (2.6)	44 (2.4)	28 (1.4)

Missing	1,042 (1.8)	386 (1.7)	60 (0.5)	19 (0.7)	518 (3.1)	46 (2.5)	13 (0.7)
Sex, n (%)							
Female	30,782 (53.3)	12,856 (56.4)	6,115 (52.6)	1,721 (59.8)	8,487 (51.0)	829 (45.1)	774 (39.5)
Male	26,939 (46.6)	9,920 (43.5)	5,515 (47.4)	1,152 (40.0)	8,160 (49.0)	1,008 (54.9)	1,184 (60.4)
Missing	27 (<0.1)	12 (<0.1)	3 (<0.1)	4 (<0.1)	6 (<0.1)	0 (0.0)	2 (<0.1)
Race, n (%)							
African American	13,183 (22.8)	3,473 (15.2)	3,178 (27.3)	790 (27.5)	4,675 (28.1)	551 (30.0)	516 (26.3)
Asian	1,848 (3.2)	639 (2.8)	438 (3.8)	100 (3.5)	555 (3.3)	41 (2.2)	75 (3.8)
Caucasian	29,074 (50.3)	13,746 (60.3)	4,653 (40.0)	1,337 (46.5)	7,538 (45.3)	849 (46.2)	951 (48.5)
Missing	13,643 (23.6)	4,930 (21.6)	3,364 (28.9)	650 (22.6)	3,885 (23.3)	396 (21.6)	418 (21.3)
Ethnicity, n (%)							
Hispanic	11,932 (20.7)	3,942 (17.3)	3,378 (29.0)	646 (22.5)	3,298 (19.8)	332 (18.1)	336 (17.1)
Non-Hispanic	38,988 (67.5)	15,485 (68.0)	7,121 (61.2)	1,987 (69.1)	11,648 (69.9)	1,294 (70.4)	1,453 (74.1)
Missing	6,828 (11.8)	3,361 (14.7)	1,134 (9.7)	244 (8.5)	1,707 (10.3)	211 (11.5)	171 (8.7)

Region, n (%)							
Midwest	22,133 (38.3)	8,137 (35.7)	5,250 (45.1)	1,364 (47.4)	5,686 (34.1)	830 (45.2)	866 (44.2)
Northwest	20,671 (35.8)	8,018 (35.2)	3,261 (28.0)	875 (30.4)	7,375 (44.3)	496 (27.0)	646 (33.0)
South	8,548 (14.8)	3,463 (15.2)	2,004 (17.2)	367 (12.8)	2,212 (13.3)	271 (14.8)	231 (11.8)
West	4,430 (7.7)	2,379 (10.4)	673 (5.8)	169 (5.9)	873 (5.2)	178 (9.7)	158 (8.1)
Missing	1,966 (3.4)	791 (3.5)	445 (3.8)	102 (3.5)	507 (3.0)	62 (3.4)	59 (3.0)
Division, n (%)							
East North Central	15,381 (26.6)	4,833 (21.2)	3,822 (32.9)	982 (34.1)	4,536 (27.2)	551 (30.0)	657 (33.5)
East South Central	1,769 (3.1)	664 (2.9)	404 (3.5)	62 (2.2)	472 (2.8)	124 (6.8)	43 (2.2)
Middle Atlantic	15,163 (26.3)	6,516 (28.6)	1,718 (14.8)	527 (18.3)	5,622 (33.8)	338 (18.4)	442 (22.6)
Mountain	2,221 (3.8)	1,281 (5.6)	257 (2.2)	48 (1.7)	457 (2.7)	78 (4.2)	100 (5.1)
New England	5,478 (9.5)	1,491 (6.5)	1,538 (13.2)	345 (12.0)	1,744 (10.5)	158 (8.6)	202 (10.3)
Other/unknown	2,085 (3.6)	847 (3.7)	467 (4.0)	112 (3.9)	532 (3.2)	64 (3.5)	63 (3.2)
Pacific	2,205 (3.8)	1,096 (4.8)	416 (3.6)	121 (4.2)	414 (2.5)	100 (5.4)	58 (3.0)

South Atlantic/ West South Central	6,767 (11.7)	2,792 (12.3)	1,599 (13.7)	303 (10.5)	1,738 (10.4)	147 (8.0)	188 (9.6)
West North Central	6,679 (11.6)	3,268 (14.3)	1,412 (12.1)	377 (13.1)	1,138 (6.8)	277 (15.1)	207 (10.6)
Smoking status, n (%)							
Current smoker	413 (0.7)	193 (0.8)	128 (1.1)	13 (0.5)	61 (0.4)	13 (0.7)	5 (0.3)
Previously smoked	740 (1.3)	417 (1.8)	111 (1.0)	27 (0.9)	145 (0.9)	25 (1.4)	15 (0.8)
Never smoked	2,831 (4.9)	1,468 (6.4)	750 (6.4)	121 (4.2)	404 (2.4)	50 (2.7)	38 (1.9)
Missing	53,764 (93.1)	20,710 (90.9)	10,644 (91.5)	2,716 (94.4)	16,043 (96.3)	1,749 (95.2)	1,902 (97.0)
Obese, n (%)							
No	47,796 (81.0)	17,883 (78.5)	10,267 (88.3)	2,297 (79.8)	13,407 (80.5)	1,431 (77.9)	1,511 (77.1)
Yes	10,952 (19.0)	4,905 (21.5)	1,366 (11.7)	580 (20.2)	3,246 (19.5)	406 (22.1)	449 (22.9)
Pregnant n (%)							
No	56,137 (97.2)	22,246 (97.6)	11,409 (98.1)	2,802 (97.4)	15,931 (95.7)	1,813 (98.7)	1,936 (98.8)
Yes	1,611 (2.8)	542 (2.4)	224 (1.9)	75 (2.6)	722 (4.3)	24 (1.3)	24.2 (1.2)
Insurance, n (%)							

Commercial	29,145 (50.5)	13,134 (57.6)	5,672 (48.8)	1,482 (51.5)	7,243 (43.5)	758 (41.3)	856 (43.7)
Medicaid	8,652 (15.0)	2,341 (10.3)	2,223 (19.1)	542 (18.8)	2,891 (17.4)	312 (17.0)	343 (17.5)
Medicare	8,774 (15.2)	3,173 (13.9)	788 (6.8)	245 (8.5)	3674 (22.1)	435 (23.7)	459 (23.4)
Other payor type	4,004 (6.9)	1,282 (5.6)	1,071 (9.2)	211 (7.3)	1,188 (7.1)	129 (7.0)	123 (6.3)
Uninsured	4,833 (8.4)	1,542 (6.8)	1,731 (14.9)	281 (9.8)	1,069 (6.4)	111 (6.0)	99 (5.1)
Missing	2,340 (4.1)	1,316 (5.8)	148 (1.3)	116 (4.0)	588 (3.5)	92 (5.0)	80 (4.1)
Month of COVID-19 diagnosis, n (%)							
Feb 2020	115 (0.2)	61 (0.3)	3 (<0.1)	4 (0.1)	34 (0.2)	5 (0.3)	8 (0.4)
Mar 2020	8,197 (14.2)	1,527 (6.7)	1,893 (16.3)	527 (18.3)	3,288 (19.7)	306 (16.7)	656 (33.5)
Apr 2020	18,591 (32.2)	6,018 (26.4)	3,480 (29.9)	926 (32.2)	6,684 (40.1)	676 (36.8)	807 (41.2)
May 2020	14,188 (24.6)	6,154 (27.0)	2,703 (23.2)	729 (25.3)	3755 (22.5)	477 (26.0)	370 (18.9)
Jun 2020	14,832 (25.7)	7,846 (34.4)	3,073 (26.4)	597 (20.8)	2,826 (17.0)	371 (20.2)	119 (6.1)
Jul 2020	1,825 (3.2)	1,182 (5.2)	481 (4.1)	94 (3.3)	66 (0.4)	2 (0.1)	0 (0.0)
Diabetes with complication, n (%)							

No	53,804 (93.2)	21,658 (95.0)	11,356 (97.6)	2,774 (96.4)	14,812 (88.9)	1,553 (84.5)	1,651 (84.2)
Yes	3,944 (6.8)	1,130 (5.0)	277 (2.4)	103 (3.6)	1,841 (11.1)	284 (15.5)	309 (15.8)
Congestive heart failure, n (%)							
No	54,048 (93.6)	21,702 (95.2)	11,430 (98.3)	2,800 (97.3)	14,938 (89.7)	1,553 (84.5)	1,625 (82.9)
Yes	3,700 (6.4)	1,086 (4.8)	203 (1.7)	77 (2.7)	1,715 (10.3)	284 (15.5)	335 (17.1)
Cerebrovascular disease, n (%)							
No	55,258 (95.7)	21,957 (96.4)	11,485 (98.7)	2,811 (97.7)	15,547 (93.4)	1,662 (90.5)	1,796 (91.6)
Yes	2,490 (4.3)	831 (3.6)	148 (1.3)	66 (2.3)	1,106 (6.6)	175 (9.5)	164 (8.4)
Moderate or severe renal disease, n (%)							
No	53,066 (91.9)	21,454 (94.1)	11,387 (97.9)	2,759 (95.9)	14,387 (86.4)	1,497 (81.5)	1,582 (80.7)
Yes	4,682 (8.1)	1,334 (5.9)	246 (2.1)	118 (4.1)	2,266 (13.6)	340 (18.5)	378 (19.3)
Diabetes without complication, n (%)							
No	47,489 (82.2)	19,807 (86.9)	10,435 (89.7)	2,522 (87.7)	12,313 (73.9)	1,184 (64.5)	1,228 (62.7)

Yes	10,259 (17.8)	2,981 (13.1)	1,198 (10.3)	355 (12.3)	4,340 (26.1)	653 (35.5)	732 (37.3)
Chronic pulmonary disease, n (%)							
No	47,794 (82.8)	19,225 (84.4)	10,203 (87.7)	2,359 (82.0)	13,145 (78.9)	1,439 (78.3)	1,423 (72.6)
Yes	9,954 (17.2)	3,563 (15.6)	1,430 (12.3)	518 (18.0)	3,508 (21.1)	398 (21.7)	537 (27.4)
Mild liver disease, n (%)							
No	55,817 (96.7)	22,095 (97.0)	11,441 (98.3)	2,788 (96.9)	15,890 (95.4)	1,746 (95.0)	1,857 (94.7)
Yes	1,931 (3.3)	693 (3.0)	192 (1.7)	89 (3.1)	763 (4.6)	91 (5.0)	103 (5.3)
Peripheral vascular disease, n (%)							
No	55,049 (95.3)	21,773 (95.5)	11,461 (98.5)	2,808 (97.6)	15,516 (93.2)	1,674 (91.1)	1,817 (92.7)
Yes	2,699 (4.7)	1,015 (4.5)	172 (1.5)	69 (2.4)	1,137 (6.8)	163 (8.9)	143 (7.3)
Cancer, n (%)							
No	53,687 (93.0)	20,822 (91.4)	11,198 (96.3)	2,686 (93.4)	15,465 (92.9)	1,689 (91.9)	1,827 (93.2)
Yes	4,061 (7.0)	1,966 (8.6)	435 (3.7)	191 (6.6)	1,188 (7.1)	148 (8.1)	133 (6.8)

Myocardial infarction, n (%)							
No	55,528 (96.2)	22,273 (97.7)	11,459 (98.5)	2,828 (98.3)	15,528 (93.2)	1,659 (90.3)	1,781 (90.9)
Yes	2,220 (3.8)	515 (2.3)	174 (1.5)	49 (1.7)	1,125 (6.8)	178 (9.7)	179 (9.1)
Dementia, n (%)							
No	55,833 (96.7)	22,213 (97.5)	11,541 (99.2)	2,843 (98.8)	15,663 (94.1)	1,697 (92.4)	1,876 (95.7)
Yes	1,915 (3.3)	575 (2.5)	92 (0.8)	34 (1.2)	990 (5.9)	140 (7.6)	84 (4.3)
Peptic ulcer disease, n (%)							
No	57,305 (99.2)	22,620 (99.3)	11,599 (99.7)	2,864 (99.5)	16,494 (99.0)	1,812 (98.6)	1,916 (97.8)
Yes	443 (0.8)	168 (0.7)	34 (0.3)	13 (0.5)	159 (1.0)	25 (1.4)	44 (2.2)
Hemiplegia, n (%)							
No	57,192 (99.0)	22,647 (99.4)	11,596 (99.7)	2,870 (99.8)	16,402 (98.5)	1,783 (97.1)	1,894 (96.6)
Yes	556 (1.0)	141 (0.6)	37 (0.3)	7 (0.2)	251 (1.5)	54 (2.9)	66 (3.4)

Rheumatics, n (%)							
No	56,674 (98.1)	22,348 (98.1)	11,515 (99.0)	2,831 (98.4)	16,285 (97.8)	1,790 (97.4)	1,905 (97.2)
Yes	1,074 (1.9)	440 (1.9)	118 (1.0)	46 (1.6)	368 (2.2)	47 (2.6)	55 (2.8)
Metastatic solid tumor, n (%)							
No	57,146 (99.0)	22,460 (98.6)	11,601 (99.7)	2,858 (99.3)	16,472 (98.9)	1,810 (98.5)	1,945 (99.2)
Yes	602 (1.0)	328 (1.4)	32 (0.3)	19 (0.7)	181 (1.1)	27 (1.5)	15 (0.8)
Moderate or severe liver disease, n (%)							
No	57,495 (99.6)	22,703 (99.6)	11,625 (99.9)	2,872 (99.8)	16,547 (99.4)	1,818 (99.0)	1,930 (98.5)
Yes	253 (0.4)	85 (0.4)	8 (0.1)	5 (0.2)	106 (0.6)	19 (1.0)	30 (1.5)
AIDS, n (%)							
No	56,640 (98.1)	22,229 (97.5)	11,500 (98.9)	2,813 (97.8)	1,6376 (98.3)	1,794 (97.7)	1,928 (98.4)
Yes	1,108 (1.9)	559 (2.5)	133 (1.1)	64 (2.2)	277 (1.7)	43 (2.3)	32 (1.6)
Mean weighted CCI (SD)	1.20 (2.06)	1.09 (2.03)	0.56 (1.32)	0.88 (1.76)	1.64 (2.31)	2.15 (2.46)	2.13 (2.39)

COPD, n (%)							
No	54,029 (93.6)	21,688 (95.2)	11,382 (97.8)	2,770 (96.3)	14,961 (89.8)	1,594 (86.8)	1,634 (83.4)
Yes	3,719 (6.4)	1,100 (4.8)	251 (2.2)	107 (3.7)	1,692 (10.2)	243 (13.2)	326 (16.6)
Diabetes, n (%)							
No	46,563 (80.6)	19,539 (85.7)	10,380 (89.2)	2,489 (86.5)	11,878 (71.3)	1,119 (60.9)	1,158 (59.1)
Yes	11,185 (19.4)	3,249 (14.3)	1,253 (10.8)	388 (13.5)	4,775 (28.7)	718 (39.1)	802 (40.9)
Hypertension, n (%)							
No	37,710 (65.3)	16,023 (70.3)	9,335 (80.2)	2,130 (74.0)	8,643 (51.9)	793 (43.2)	786 (40.1)
Yes	20,038 (34.7)	6,765 (29.7)	2,298 (19.8)	747 (26.0)	8,010 (48.1)	1,044 (56.8)	1,174 (59.9)
Asthma, n (%)							
No	51,726 (89.6)	20,663 (90.7)	10,587 (91.0)	2,517 (87.5)	14,604 (87.7)	1,652 (89.9)	1,703 (86.9)
Yes	6,022 (10.4)	2,125 (9.3)	1,046 (9.0)	360 (12.5)	2,049 (12.3)	185 (10.1)	257 (13.1)
Chronic renal disease, n (%)							

No	53,421 (92.5)	21,585 (94.7)	11,411 (98.1)	2,768 (96.2)	14,520 (87.2)	1,524 (83.0)	1,613 (82.3)
Yes	4,327 (7.5)	1,203 (5.3)	222 (1.9)	109 (3.8)	2,133 (12.8)	313 (17.0)	347 (17.7)
Other chronic respiratory disease, n (%)							
No	55,858 (96.7)	21,900 (96.1)	11,365 (97.7)	2,738 (95.2)	16,186 (97.2)	1,779 (96.8)	1,890 (96.4)
Yes	1,890 (3.3)	888 (3.9)	268 (2.3)	139 (4.8)	467 (2.8)	58 (3.2)	70 (3.6)
Chronic ischemic heart disease, n (%)							
No	52,308 (90.6)	20,978 (92.1)	11,270 (96.9)	2,733 (95.0)	14,236 (85.5)	1,491 (81.2)	1,600 (81.6)
Yes	5,440 (9.4)	1,810 (7.9)	363 (3.1)	144 (5.0)	2,417 (14.5)	346 (18.8)	360 (18.4)
End stage renal disease, n (%)							
No	56,530 (97.9)	22,463 (98.6)	11,560 (99.4)	2,849 (99.0)	16,079 (96.6)	1,740 (94.7)	1,839 (93.8)
Yes	1,218 (2.1)	325 (1.4)	73 (0.6)	28 (1.0)	574 (3.4)	97 (5.3)	121 (6.2)
Liver disease							
No	55,348 (95.8)	21,961 (96.4)	11,407 (98.1)	2,775 (96.5)	15,683 (94.2)	1,718 (93.5)	1,804 (92.0)

Yes	2,400 (4.2)	827 (3.6)	226 (1.9)	102 (3.5)	970 (5.8)	119 (6.5)	156 (8.0)
HIV, n (%)							
No	57,000 (98.7)	22,425 (98.4)	11,533 (99.1)	2,836 (98.6)	16,456 (98.8)	1,811 (98.6)	1,939 (98.9)
Yes	748 (1.3)	363 (1.6)	100 (0.9)	41 (1.4)	197 (1.2)	26 (1.4)	21 (1.1)
Immunocompromised, n (%)							
No	53,530 (92.7)	21,906 (96.1)	11,403 (98.0)	2,802 (97.4)	14,429 (86.6)	1,490 (81.1)	1,500 (76.5)
Yes	4,218 (7.3)	882 (3.9)	230 (2.0)	75 (2.6)	2,224 (13.4)	347 (18.9)	460 (23.5)

AIDS, acquired immunodeficiency syndrome; CCI, Charlson Comorbidity Index; COPD, chronic obstructive pulmonary disease; COVID-19, coronavirus disease 2019; ER, emergency room; HIV, human immunodeficiency virus; ICU, intensive care unit; SD, standard deviation

Supplemental Table 4 Full list of long-term outcomes that occurred >30–≤90 days or >90–≤180 days post COVID-19 diagnosis or hospitalization

	1. Outpatient (N=22,788)		2. ER on diagnosis date (N=11,633)		3. ER (N=2,877)		4. Hospitalization without ICU (N=16,653)		5. ICU without ventilation (N=1,837)		6. ICU with ventilation (N=1,960)	
	>30–≤90 days	**>90–≤180 days**	**>30–≤90 days**	**>90–≤180 days**	**>30–≤90 days**	**>90–≤180 days**	**>30–≤90 days**	**>90–≤180 days**	**>30–≤90 days**	**>90–≤180 days**	**>30–≤90 days**	**>90–≤180 days**
Pneumonia, n (%)												
No	22,368 (98.2)	22,582 (99.1)	11,520 (99.0)	11,579 (99.5)	2,804 (97.5)	2,848 (99.0)	15,612 (93.7)	16,187 (97.2)	1,689 (91.9)	1,768 (96.2)	1,687 (86.1)	1,801 (91.9)
Yes	420 (1.8)	206 (0.9)	113 (1.0)	54 (0.5)	73 (2.5)	29 (1.0)	1,041 (6.3)	466 (2.8)	148 (8.1)	69 (3.8)	273 (13.9)	159 (8.1)
Asthma, n (%)												
No	22,487 (98.7)	22,459 (98.6)	11,532 (99.1)	11,503 (98.9)	2,825 (98.2)	2,822 (98.1)	16,424 (98.6)	16,410 (98.5)	1,810 (98.5)	1,815 (98.8)	1,919 (97.9)	1,922 (98.1)
Yes	301 (1.3)	329 (1.4)	101 (0.9)	130 (1.1)	52 (1.8)	55 (1.9)	229 (1.4)	243 (1.5)	27 (1.5)	22 (1.2)	41 (2.1)	38 (1.9)
COPD, n (%)												

No	22,626 (99.3)	22,776 (99.9)	11,615 (99.8)	11,606 (99.8)	2,865 (99.6)	2,858 (99.3)	16,460 (98.8)	16,442 (98.7)	1,802 (98.1)	1,804 (98.2)	1,894 (96.6)	1,919 (97.9)
Yes	162 (0.7)	179 (0.8)	18 (0.2)	28 (0.2)	12 (0.4)	19 (0.7)	193 (1.2)	211 (1.3)	35 (1.9)	33 (1.8)	66 (3.4)	41 (2.1)
Influenza, n (%)												
No	22,783 (100.0)	22,776 (99.9)	11,630 (100.0)	11,631 (100.0)	2,875 (99.9)	2,877 (100.0)	16,646 (100.0)	16,648 (100.0)	1,837 (100.0)	1,837 (100.0)	1,960 (100.0)	1,960 (100.0)
Yes	5 (0.0)	12 (0.1)	3 (0.0)	2 (0.0)	2 (0.1)	0 (0.0)	7 (0.0)	5 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Stroke, n (%)												
No	22,695 (99.6)	22,696 (99.6)	11,619 (99.9)	11,611 (99.8)	2,865 (99.6)	2,865 (99.6)	16,535 (99.3)	16,506 (99.1)	1,813 (98.7)	1,811 (98.6)	1,935 (98.7)	1,926 (98.3)
Yes	93 (0.4)	92 (0.4)	14 (0.1)	22 (0.2)	12 (0.4)	12 (0.4)	118 (0.7)	147 (0.9)	24 (1.3)	26 (1.4)	25 (1.3)	34 (1.7)
Anxiety, n (%)												
No	22,250 (97.6)	22,169 (97.3)	11,491 (98.8)	11,451 (98.4)	2,779 (96.6)	2,774 (96.4)	16,274 (97.7)	16,268 (97.7)	1,793 (97.6)	1,794 (97.7)	1,889 (96.4)	1,896 (96.7)
Yes	538 (2.4)	619 (2.7)	142 (1.2)	182 (1.6)	98 (3.4)	103 (3.6)	379 (2.3)	385 (2.3)	44 (2.4)	43 (2.3)	71 (3.6)	64 (3.3)

Depression, n (%)												
No	22,456 (98.5)	22,361 (98.1)	11,535 (99.2)	11,502 (98.9)	2,833 (98.5)	2,822 (98.1)	16,375 (98.3)	16,314 (98.0)	1,789 (97.4)	1,786 (97.2)	1,907 (97.3)	1,908 (97.3)
Yes	332 (1.5)	427 (1.9)	98 (0.8)	131 (1.1)	44 (1.5)	55 (1.9)	278 (1.7)	339 (2.0)	48 (2.6)	51 (2.8)	53 (2.7)	52 (2.7)
Myocardial infarction, n (%)												
No	22,691 (99.6)	22,671 (99.5)	11,617 (99.9)	11,617 (99.9)	2,866 (99.6)	2,868 (99.7)	16,497 (99.1)	16,492 (99.0)	1,810 (98.5)	1,806 (98.3)	1,927 (98.3)	1,926 (98.3)
Yes	97 (0.4)	117 (0.5)	16 (0.1)	16 (0.1)	11 (0.4)	9 (0.3)	156 (0.9)	161 (1.0)	27 (1.5)	31 (1.7)	33 (1.7)	34 (1.7)
Interstitial lung disease (fibrosis), n (%)												
No	22,741 (99.8)	22,728 (99.7)	11,623 (99.9)	11,621 (99.9)	2,874 (99.9)	2,873 (99.9)	16,592 (99.6)	16,578 (99.5)	1,830 (99.6)	1,828 (99.5)	1,929 (98.4)	1,922 (98.1)
Yes	47 (0.2)	60 (0.3)	10 (0.1)	12 (0.1)	3 (0.1)	4 (0.1)	61 (0.4)	75 (0.5)	7 (0.4)	9 (0.5)	31 (1.6)	38 (1.9)
Dsypnea, n (%)												
No	21,660 (95.1)	21,567 (94.6)	11,311 (97.2)	11,329 (97.4)	2,675 (93.0)	2,649 (92.1)	15,781 (94.8)	15,838 (95.1)	1,720 (93.6)	1,723 (93.8)	1,759 (89.7)	1,783 (91.0)

Yes	1,128 (4.9)	1,221 (5.4)	322 (2.8)	304 (2.6)	202 (7.0)	228 (7.9)	872 (5.2)	815 (4.9)	117 (6.4)	114 (6.2)	201 (10.3)	177 (9.0)
Respiratory failure, n (%)												
No	22,614 (99.2)	22,654 (99.4)	11,606 (99.8)	11,609 (99.8)	2,863 (99.5)	2,868 (99.7)	16,199 (97.3)	16,380 (98.4)	1,757 (95.6)	1,780 (96.9)	1,685 (86.0)	1,801 (91.9)
Yes	174 (0.8)	134 (0.6)	27 (0.2)	24 (0.2)	14 (0.5)	9 (0.3)	454 (2.7)	273 (1.6)	80 (4.4)	57 (3.1)	275 (14.0)	159 (8.1)
Pulmonary hypertension, n (%)												
No	22,719 (99.7)	22,716 (99.7)	11,626 (99.9)	11,622 (99.9)	2,872 (99.8)	2,874 (99.9)	16,566 (99.5)	16,551 (99.4)	1,824 (99.3)	1,821 (99.1)	1,944 (99.2)	1,927 (98.3)
Yes	69 (0.3)	72 (0.3)	7 (0.1)	11 (0.1)	5 (0.2)	3 (0.1)	87 (0.5)	102 (0.6)	13 (0.7)	16 (0.9)	16 (0.8)	33 (1.7)
Pulmonary embolism, n (%)												
No	22,714 (99.7)	22,719 (99.7)	11,622 (99.9)	11,623 (99.9)	2,865 (99.6)	2,863 (99.5)	16,478 (98.9)	16,513 (99.2)	1,809 (98.5)	1,815 (98.8)	1,918 (97.9)	1,938 (98.9)
Yes	74 (0.3)	69 (0.3)	11 (0.1)	10 (0.1)	12 (0.4)	14 (0.5)	175 (1.1)	140 (0.8)	28 (1.5)	22 (1.2)	42 (2.1)	22 (1.1)
Bronchitis, n (%)												

No	22,707 (99.6)	22,705 (99.6)	11,611 (99.8)	11,619 (99.9)	2,862 (99.5)	2,868 (99.7)	16,583 (99.6)	16,598 (99.7)	1,830 (99.6)	1,832 (99.7)	1,939 (98.9)	1,940 (99.0)
Yes	81 (0.4)	83 (0.4)	22 (0.2)	14 (0.1)	15 (0.5)	9 (0.3)	70 (0.4)	55 (0.3)	7 (0.4)	5 (0.3)	21 (1.1)	20 (1.0)
Emphysema, n (%)												
No	22,727 (99.7)	22,722 (99.7)	11,626 (99.9)	11,620 (99.9)	2,872 (99.8)	2,870 (99.8)	16,591 (99.6)	16,577 (99.5)	1,815 (98.8)	1,822 (99.2)	1,941 (99.0)	1,944 (99.2)
Yes	61 (0.3)	66 (0.3)	7 (0.1)	13 (0.1)	5 (0.2)	7 (0.2)	62 (0.4)	76 (0.5)	22 (1.2)	15 (0.8)	19 (1.0)	16 (0.8)
Bronchiectasis, n (%)												
No	22,765 (99.9)	22,763 (99.9)	11,632 (100.0)	11,629 (100.0)	2,876 (100.0)	2,874 (99.9)	16,630 (99.9)	16,625 (99.8)	1,836 (99.9)	1,836 (99.9)	1,951 (99.5)	1,952 (99.6)
Yes	23 (0.1)	25 (0.1)	1 (0.0)	4 (0.0)	1 (0.0)	3 (0.1)	23 (0.1)	28 (0.2)	1 (0.1)	1 (0.1)	9 (0.5)	8 (0.4)
Encephalopathy, n (%)												
No	22,709 (99.7)	22,732 (99.8)	11,624 (99.9)	11,627 (99.9)	2,872 (99.8)	2,874 (99.9)	16,545 (99.4)	16,554 (99.4)	1,809 (98.5)	1,816 (98.9)	1,911 (97.5)	1,923 (98.1)
Yes	79 (0.3)	56 (0.2)	9 (0.1)	6 (0.1)	5 (0.2)	3 (0.1)	108 (0.6)	99 (0.6)	28 (1.5)	21 (1.1)	49 (2.5)	37 (1.9)

Memory loss, n (%)												
No	22,752 (99.8)	22,716 (99.7)	11,622 (99.9)	11,621 (99.9)	2,872 (99.8)	2,870 (99.8)	16,626 (99.8)	16,599 (99.7)	1,833 (99.8)	1,831 (99.7)	1,951 (99.5)	1,946 (99.3)
Yes	36 (0.2)	72 (0.3)	11 (0.1)	12 (0.1)	5 (0.2)	7 (0.2)	27 (0.2)	54 (0.3)	4 (0.2)	6 (0.3)	9 (0.5)	14 (0.7)
Confusion or disorientation, n (%)												
No	22,699 (99.6)	22,706 (99.6)	11,621 (99.9)	11,617 (99.9)	2,869 (99.7)	2,869 (99.7)	16,531 (99.3)	16,526 (99.2)	1,817 (98.9)	1,817 (98.9)	1,929 (98.4)	1,939 (98.9)
Yes	89 (0.4)	82 (0.4)	12 (0.1)	16 (0.1)	8 (0.3)	8 (0.3)	122 (0.7)	127 (0.8)	20 (1.1)	20 (1.1)	31 (1.6)	21 (1.1)
Dementia, n (%)												
No	22,694 (99.6)	22,709 (99.7)	11,628 (100.0)	11,625 (99.9)	2,870 (99.8)	2,872 (99.8)	16,494 (99.0)	16,494 (99.0)	1,810 (98.5)	1,816 (98.9)	1,944 (99.2)	1,947 (99.3)
Yes	94 (0.4)	79 (0.3)	5 (0.0)	8 (0.1)	7 (0.2)	5 (0.2)	159 (1.0)	159 (1.0)	27 (1.5)	21 (1.1)	16 (0.8)	13 (0.7)
Cardiac arrhythmia, n (%)												
No	22,627 (99.3)	22,598 (99.2)	11,594 (99.7)	11,593 (99.7)	2,860 (99.4)	2,850 (99.1)	16,515 (99.2)	16,488 (99.0)	1,819 (99.0)	1,816 (98.9)	1,935 (98.7)	1,931 (98.5)

Yes	161 (0.7)	190 (0.8)	39 (0.3)	40 (0.3)	17 (0.6)	27 (0.9)	138 (0.8)	165 (1.0)	18 (1.0)	21 (1.1)	25 (1.3)	29 (1.5)
Respiratory, n (%)												
No	20,942 (91.9)	20,942 (91.9)	11,113 (95.5)	11,149 (95.8)	2,555 (88.8)	2,566 (89.2)	14,529 (87.2)	15,054 (90.4)	1,541 (83.9)	1,615 (87.9)	1,413 (72.1)	1,579 (80.6)
Yes	1,846 (8.1)	1,846 (8.1)	520 (4.5)	484 (4.2)	322 (11.2)	311 (10.8)	2,124 (12.8)	1,599 (9.6)	296 (16.1)	222 (12.1)	547 (27.9)	381 (19.4)
CV, n (%)												
No	22,349 (98.1)	22,313 (97.9)	11,522 (99.3)	11,541 (99.2)	2,823 (98.1)	2,820 (98.0)	16,068 (96.5)	16,035 (96.3)	1,735 (94.4)	1,738 (94.6)	1,832 (93.5)	1,828 (93.3)
Yes	439 (1.9)	475 (2.1)	81 (0.7)	92 (0.8)	54 (1.9)	57 (2.0)	585 (3.5)	618 (3.7)	102 (5.6)	99 (5.4)	128 (6.5)	132 (6.7)
Mental health, n (%)												
No	21,848 (95.9)	21,707 (95.3)	11,402 (98.0)	11,340 (97.5)	2,733 (95.0)	2,723 (94.6)	15,834 (95.1)	15,787 (94.8)	1,718 (93.5)	1,719 (93.6)	1,793 (91.5)	1,816 (92.7)
Yes	940 (4.1)	1,081 (4.7)	231 (2.0)	293 (2.5)	144 (5.0)	154 (5.4)	819 (4.9)	866 (5.2)	119 (6.5)	118 (6.4)	167 (8.5)	144 (7.3)
Cancer, n (%)												

No	22,561 (99.4)	22,637 (99.3)	11,609 (99.8)	11,603 (99.7)	2,859 (99.4)	2,863 (99.5)	16,558 (99.4)	16,547 (99.4)	1,828 (99.5)	1,825 (99.3)	1,938 (98.9)	1,946 (99.3)
Yes	137 (0.6)	151 (0.7)	24 (0.2)	30 (0.3)	18 (0.6)	14 (0.5)	95 (0.6)	106 (0.6)	9 (0.5)	12 (0.7)	22 (1.1)	14 (0.7)
Respiratory and CV, n (%)												
No	22,657 (99.4)	22,664 (99.5)	11,617 (99.9)	11,603 (99.7)	2,860 (99.4)	2,858 (99.3)	16,452 (98.8)	16,472 (98.9)	1,807 (98.8)	1,811 (98.6)	1,897 (96.8)	1,913 (97.6)
Yes	131 (0.6)	124 (0.5)	16 (0.1)	30 (0.3)	17 (0.6)	19 (0.7)	201 (1.2)	181 (1.1)	30 (1.6)	26 (1.4)	63 (3.2)	47 (2.4)
Respiratory and mental health, n (%)												
No	22,583 (99.1)	22,572 (99.1)	11,568 (99.4)	11,570 (99.5)	2,829 (98.3)	2,844 (98.9)	16,391 (98.4)	16,408 (98.5)	1,792 (97.6)	1,803 (98.1)	1,877 (95.8)	1,910 (97.4)
Yes	205 (0.9)	216 (0.9)	65 (0.6)	63 (0.5)	48 (1.7)	33 (1.1)	262 (1.6)	245 (1.5)	45 (2.4)	34 (1.9)	83 (4.2)	50 (2.6)
Mental health and CV, n (%)												
No	22,735 (99.8)	22,739 (99.8)	11,625 (99.9)	11,624 (99.9)	2,873 (99.9)	2,875 (99.9)	16,596 (99.7)	16,589 (99.6)	1,827 (99.5)	1,826 (99.4)	1,953 (99.6)	1,949 (99.4)
Yes	53 (0.2)	49 (0.2)	8 (0.1)	9 (0.1)	4 (0.1)	2 (0.1)	57 (0.3)	64 (0.4)	10 (0.5)	11 (0.6)	7 (0.4)	11 (0.6)

Respiratory, CV, and mental health, n (%)												
No	22,731 (99.7)	22,736 (99.8)	11,624 (99.9)	11,623 (99.9)	2,871 (99.8)	2,868 (99.7)	16,555 (99.4)	16,569 (99.5)	1,820 (99.1)	1,822 (99.2)	1,930 (98.5)	1,929 (98.4)
Yes	57 (0.3)	52 (0.2)	9 (0.1)	10 (0.1)	6 (0.2)	9 (0.3)	98 (0.6)	84 (0.5)	17 (0.9)	15 (0.8)	30 (1.5)	31 (1.6)
No conditions, n (%)												
No	2,722 (11.9)	2,909 (12.8)	725 (6.2)	747 (6.4)	439 (15.3)	450 (15.6)	2,812 (16.9)	2,425 (14.6)	398 (21.7)	338 (18.4)	629 (32.1)	487 (24.8)
Yes	20,066 (88.1)	19,879 (87.2)	10,908 (93.8)	10,886 (93.6)	2,438 (84.7)	2,427 (84.4)	13,841 (83.1)	14,228 (85.4)	1,439 (78.3)	1,499 (81.6)	1,331 (67.9)	1,473 (75.2)

COPD, chronic obstructive pulmonary disease; COVID-19, coronavirus disease 2019; CV, cardiovascular; ER, emergency room; ICU, intensive care unit

Supplemental Table 5 Full list of risk ratios and 95% confidence intervals of covariates associated with the occurrence of new respiratory, cardiovascular, and mental health conditions at >30–≤90 days or >90–≤180 days post COVID-19 diagnosis or hospital discharge

RR (95% CI)	Respiratory conditions		Cardiovascular conditions		Mental health conditions	
	>30–≤90 days	>90–≤180 days	>30–≤90 days	>90–≤180 days	>30–≤90 days	>90–≤180 days
Age group						
18–29 years	1.18 (0.80, 1.79)	1.08 (0.76, 1.56)	NA*	3.18 (0.94, 19.53)	2.67 (1.59, 4.85)	1.81 (1.22, 2.79)
30–39 years	1.90 (1.32, 2.82)	1.59 (1.14, 2.27)	NA*	6.09 (1.88, 36.34)	3.00 (1.79, 5.43)	1.87 (1.26, 2.88)
40–49 years	2.44 (1.71, 3.58)	1.96 (1.41, 2.77)	NA*	6.68 (2.05, 39.85)	2.73 (1.63, 4.96)	1.91 (1.29, 2.94)
50–64 years	2.84 (2.01, 4.13)	1.92 (1.40, 2.71)	NA*	8.03 (2.52, 47.28)	2.94 (1.77, 5.28)	1.63 (1.11, 2.50)
65–74 years	2.68 (1.86, 3.95)	1.88 (1.34, 2.70)	NA*	10.46 (3.10, 62.22)	2.30 (1.35, 4.22)	1.28 (0.85, 2.02)
75–84 years	2.43 (1.65, 3.65)	1.52 (1.04, 2.24)	NA*	18.40 (5.23, 107.58)	2.61 (1.50, 4.85)	1.29 (0.83, 2.07)
≥85 years	1.80 (1.10, 2.95)	1.62 (1.01, 2.58)	NA*	15.49 (2.47, 111.25)	3.05 (1.59, 6.02)	1.73 (1.00, 2.98)
Sex						
Male	0.91 (0.84, 0.99)	0.84 (0.77, 0.92)	1.18 (0.92, 1.51)	1.26 (0.98, 1.61)	0.63 (0.56, 0.70)	0.53 (0.47, 0.59)
Race						

Caucasian	1.02 (0.94, 1.12)	1.13 (1.02, 1.24)	1.10 (0.83, 1.47)	1.10 (0.83, 1.48)	1.48 (1.31, 1.69)	1.792 (1.59, 2.03)
Asian	0.99 (0.81, 1.19)	1.00 (0.80, 1.24)	0.72 (0.33, 1.36)	0.84 (0.41, 1.56)	0.74 (0.50, 1.04)	0.81 (0.56, 1.12)
Ethnicity						
Non-Hispanic	1.12 (0.97, 1.29)	1.38 (1.17, 1.62)	2.56 (1.53, 4.63)	2.12 (1.30, 3.70)	1.40 (1.14, 1.73)	1.54 (1.27, 1.87)
Obesity	1.33 (1.21, 1.46)	1.38 (1.24, 1.53)	1.08 (0.76, 1.51)	1.18 (0.83, 1.63)	1.41 (1.24, 1.59)	1.49 (1.32, 1.67)
Insurance						
Medicaid	1.01 (0.89, 1.15)	1.07 (0.94, 1.22)	1.24 (0.85, 1.76)	1.23 (0.84, 1.75)	1.39 (1.18, 1.64)	1.41 (1.21, 1.63)
Medicare	1.06 (0.93, 1.21)	1.06 (0.91, 1.22)	1.45 (0.94, 2.21)	1.51 (1.00, 2.26)	1.53 (1.29, 1.82)	1.46 (1.23, 1.72)
Other payor type	0.77 (0.64, 0.93)	1.07 (0.90, 1.27)	1.00 (0.56, 1.65)	1.10 (0.64, 1.77)	0.97 (0.75, 1.23)	1.03 (0.82, 1.28)
Uninsured	0.71 (0.58, 0.85)	0.65 (0.52, 0.80)	0.96 (0.55, 1.57)	0.66 (0.33, 1.16)	0.76 (0.56, 0.99)	0.63 (0.47, 0.82)
Sub-cohort						
ER on diagnosis	0.64 (0.56, 0.74)	0.56 (0.48, 0.65)	0.45 (0.27, 0.71)	0.59 (0.38, 0.89)	0.60 (0.49, 0.72)	0.65 (0.55, 0.78)
ER	1.39 (1.17, 1.65)	1.33 (1.10, 1.58)	1.41 (0.83, 2.27)	1.57 (0.95, 2.47)	1.22 (0.95, 1.54)	1.20 (0.95, 1.50)
Hospitalisation without ICU	1.33 (1.20, 1.47)	1.02 (0.91, 1.14)	1.74 (1.28, 2.35)	1.42 (1.04, 1.94)	1.03 (0.89, 1.18)	1.08 (0.95, 1.23)
ICU without ventilation	1.69 (1.39, 2.03)	1.18 (0.93, 1.47)	1.69 (0.75, 3.28)	2.41 (1.25, 4.23)	1.31 (0.99, 1.71)	1.34 (1.02, 1.73)

ICU with ventilation	2.64 (2.27, 3.04)	1.86 (1.55, 2.21)	3.16 (1.83, 5.18)	2.65 (1.49, 4.43)	1.89 (1.51, 2.35)	1.52 (1.20, 1.91)
Month of COVID-19 diagnosis						
Feb–Apr 2020	1.08 (0.97, 1.20)	1.06 (0.93, 1.19)	0.88 (0.62, 1.24)	0.96 (0.66, 1.38)	0.85 (0.72, 1.01)	0.93 (0.80, 1.09)
May 2020	0.75 (0.67, 0.83)	0.92 (0.82, 1.04)	0.76 (0.55, 1.06)	1.21 (0.89, 1.65)	0.82 (0.71, 0.95)	0.90 (0.79, 1.03)
Jun 2020	0.58 (0.51, 0.65)	0.72 (0.63, 0.81)	0.68 (0.48, 0.96)	0.75 (0.52, 1.07)	0.72 (0.62, 0.84)	0.74 (0.64, 0.86)
Jul 2020	0.48 (0.34, 0.65)	0.61 (0.45, 0.82)	0.33 (0.08, 0.88)	0.81 (0.34, 1.65)	0.60 (0.40, 0.87)	0.79 (0.57, 1.08)
Weighted CCI	1.05 (1.02, 1.07)	1.07 (1.04, 1.09)	1.17 (1.10, 1.25)	1.16 (1.08, 1.23)	1.12 (1.09, 1.15)	1.14 (1.11, 1.17)

Grey and blue shading denote increased and decreased risk of a new condition occurring, respectively. Reference groups are: <18 years (age), female (sex), African American (race), Hispanic (ethnicity), non-obese (obesity), diagnosis in February and March 2020 (diagnosis month); commercial (insurance), outpatient (sub-cohort).

*Values were not calculable as the reference group (<18 years) had no new diagnoses of clinical conditions

CCI, Charlson Comorbidity Index; CI, confidence interval; COVID-19, coronavirus disease 2019; ER, emergency room; ICU, intensive care unit;

NA, not applicable; RR, risk ratio

Supplemental Table 6 Full list of risk ratios and 95% confidence intervals of covariates associated with a new cancer diagnosis at >30–≤90 days or >90–≤180 days post COVID-19 diagnosis or hospital discharge

	Risk ratio (95% CI)	
	>30–≤90 days	**>90–≤180 days**
Age group		
18–29 years	0.95 (0.31, 4.16)	1.17 (0.19, 22.13)
30–39 years	0.89 (0.29, 3.90)	2.63 (0.51, 46.94)
40–49 years	1.43 (0.49, 6.05)	3.38 (0.69, 59.46)
50–64 years	2.44 (0.90, 9.93)	8.35 (1.84, 138.73)
65–74 years	3.19 (1.13, 13.21)	13.50 (2.91, 217.30)
75–84 years	2.60 (0.86, 11.12)	15.50 (3.25, 247.65)
≥85 years	2.12 (0.53, 10.33)	8.45 (1.39, 151.23)
Sex		
Male	0.97 (0.75, 1.27)	1.03 (0.79, 1.33)
Race		
Caucasian	1.51 (1.10, 2.11)	1.05 (0.78, 1.43)
Asian	1.45 (0.66, 2.80)	1.15 (0.53, 2.20)
Ethnicity		
Non-Hispanic	2.49 (1.33, 5.28)	1.34 (0.79, 2.50)

Obesity	1.34 (0.98, 1.80)	1.34 (1.00, 1.79)
Insurance		
Medicaid	0.88 (0.52, 1.41)	0.68 (0.38, 1.13)
Medicare	1.43 (0.99, 2.07)	1.10 (0.78, 1.55)
Other payor type	0.72 (0.32, 1.39)	0.69 (0.32, 1.29)
Uninsured	1.32 (0.68, 2.32)	0.49 (0.17, 1.09)
Sub-cohort		
ER on diagnosis	0.47 (0.27, 0.76)	0.53 (0.32, 0.83)
ER	1.23 (0.68, 2.06)	0.69 (0.32, 1.30)
Hospitalization without ICU	0.71 (0.50, 0.99)	0.65 (0.47, 0.90)
ICU without ventilation	0.45 (0.16, 1.01)	0.50 (0.21, 1.02)
ICU with ventilation	1.24 (0.71, 2.06)	0.56 (0.27, 1.04)
Month of COVID-19 diagnosis		
Feb–Apr 2020	1.44 (0.98, 2.11)	1.72 (1.17, 2.51)
May 2020	1.21 (0.85, 1.72)	1.18 (0.82, 1.69)
Jun 2020	0.90 (0.61, 1.32)	1.28 (0.89, 1.84)
Jul 2020	0.70 (0.21, 1.71)	0.80 (0.24, 1.96)
Weighted CCI	1.04 (0.97, 1.11)	1.06 (0.99, 1.12)

Reference groups are: <18 years (age), female (sex), African American (race), Hispanic (ethnicity), non-obese (obesity); diagnosis in February and March 2020 (diagnosis month), commercial (insurance), outpatient (sub-cohort).

CCI, Charlson Comorbidity Index; CI, confidence interval;
COVID-19, coronavirus disease 2019; ER, emergency room; ICU, intensive care unit

The RECORD statement – checklist of items, extended from the STROBE statement, that should be reported in observational studies using routinely collected health data.

	Item No.	STROBE items	Location in manuscript where items are reported	RECORD items	Location in manuscript where items are reported
Title and abstract					
	1	(a) Indicate the study's design with a commonly used term in the title or the abstract (b) Provide in the abstract an informative and balanced summary of what was done and what was found	*(a) confirmed (Design section) (b) confirmed adequately covered in abstract*	RECORD 1.1: The type of data used should be specified in the title or abstract. When possible, the name of the databases used should be included. RECORD 1.2: If applicable, the geographic region and timeframe within which the study took place should be reported in the title or abstract. RECORD 1.3: If linkage between databases was conducted for the study, this should be clearly stated in the title or abstract.	*Abstract (Objective and Design)* *Abstract (Setting and Participants)* *Not applicable*
Introduction					
Background rationale	2	Explain the scientific background and rationale for the investigation being reported	*Confirmed present in introduction*		
Objectives	3	State specific objectives, including any prespecified hypotheses	*Specific objective stated (last paragraph of introduction; there were no pre-*		

			specified hypotheses)		
Methods					
Study Design	4	Present key elements of study design early in the paper	*Included in methods ('Patients and study design') and described in Figure 1*		
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	*Included in methods ('Database' & 'Patients and study design' sections)*		

Participants	6	*(a) Cohort study*- Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*- Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*- Give the eligibility criteria, and the sources and methods of selection of participants *(b) Cohort study*- For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*- For matched studies, give matching criteria and the number of controls per case	*(a) confirmed included in methods ('Patients and study design' section (b) not relevant (not a matched study)*	RECORD 6.1: The methods of study population selection (such as codes or algorithms used to identify subjects) should be listed in detail. If this is not possible, an explanation should be provided. RECORD 6.2: Any validation studies of the codes or algorithms used to select the population should be referenced. If validation was conducted for this study and not published elsewhere, detailed methods and results should be provided. RECORD 6.3: If the study involved linkage of databases, consider use of a flow diagram or other graphical display to demonstrate the data linkage process, including the number of individuals with linked data at each stage.	*Confirmed in methods ('Patients and study design')* *The algorithms have been used previously and is cited in the methods (Chawla et al., 2021)* *Not applicable*
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable.	*All definitions are presented in the methods ('Patients and study design', 'Modelling and*	RECORD 7.1: A complete list of codes and algorithms used to classify exposures, outcomes, confounders, and effect modifiers should be provided. If these cannot be reported, an explanation should be provided.	*Relevant lists are provided throughout the manuscript (e.g. ICD-10 codes in*

			statistical analysis', and 'Sensitivity analysis' sections)		*supplemental Tables 1 and 2, list of confounders in methods section 'Modeling and statistical analysis')*
Data sources/ measurement	8	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	*Source of data is the Optum Electronic Medical Record data, and are routinely collected by practicing physicians (detailed in methods section)*		
Bias	9	Describe any efforts to address potential sources of bias	*A sensitivity analysis was performed, and relevant controls (non-hospital setting covariates) were included in our statistical models*		
Study size	10	Explain how the study size was arrived at	*All eligible patients in the Optum dataset were included, without a prespecified study size (explained in the*		

			database and patients and study design section)		
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why	*Described in Methods section 'Modeling and statistical analysis'*		
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding (b) Describe any methods used to examine subgroups and interactions (c) Explain how missing data were addressed (d) *Cohort study*- If applicable, explain how loss to follow-up was addressed *Case-control study*- If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*- If applicable, describe analytical methods taking account of sampling strategy (e) Describe any sensitivity analyses	*a) Methods ('Modeling and statistical analysis')* *b) We do not conduct sub-group analysis* *c) Explained in discussion section* *d) We have conducted a retrospective cohort study. Regarding the loss-to follow-up, since we are not assessing the effect of a treatment, rather looking at disease severity, we assume it is non-differential.* *e) as described in methods ('Sensitivity analysis')*		

Data access and cleaning methods		..		RECORD 12.1: Authors should describe the extent to which the investigators had access to the database population used to create the study population.	*Authors had access to deidentified EMR data*
				RECORD 12.2: Authors should provide information on the data cleaning methods used in the study.	*Data cleaning methods have been described previously; the reference is cited in the Methods 'Database' section (Chawla et al).*
Linkage		..		RECORD 12.3: State whether the study included person-level, institutional-level, or other data linkage across two or more databases. The methods of linkage and methods of linkage quality evaluation should be provided.	*EMR data from hospital networks were used to form the Optum dataset. Linkage of EMR data and methods are described on Optum's website: https://www.optum.com/business/solutions/life-sciences/real-world-data/ehr-data.html*

Results					
Participants	13	(a) Report the numbers of individuals at each stage of the study (*e.g.*, numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed) (b) Give reasons for non-participation at each stage. (c) Consider use of a flow diagram	*The number of patients in the dataset and those with a COVID-19 diagnosis is given in the Methods 'Database' section. The criteria on how the population is selected is made clear in methods 'Patients and study design' section (inclusion and exclusion criteria).*	RECORD 13.1: Describe in detail the selection of the persons included in the study (*i.e.,* study population selection) including filtering based on data quality, data availability and linkage. The selection of included persons can be described in the text and/or by means of the study flow diagram.	*The criteria on how the population is selected is made clear in methods 'Patients and study design' section (inclusion and exclusion criteria)*
Descriptive data	14	(a) Give characteristics of study participants (*e.g.*, demographic, clinical, social) and information on exposures and potential confounders (b) Indicate the number of participants with missing data for each variable of interest (c) *Cohort study*- summarise follow-up time (*e.g.*, average and total amount)	*Shown in detail in tables e.g. Table 1 and also in appendix (table 1 and 2)*		
Outcome data	15	*Cohort study*- Report numbers of outcome events or summary measures over time *Case-control study*-	*Outcome data are presented in Table 2*		

		Report numbers in each exposure			
		category, or summary measures of exposure *Cross-sectional study*- Report numbers of outcome events or summary measures			
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder- adjusted estimates and their precision (e.g., 95% confidence interval). Make clear which confounders were adjusted for and why they were included (b) Report category boundaries when continuous variables were categorized (c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	*Confounder-adjusted estimates are provided; however, unadjusted results could be derived from Table 2, where the raw counts and percentages can be used to calculate raw measures of effect. Confounders we control for are described in the modelling and statistical analysis section*		
Other analyses	17	Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses	*Sensitivity analysis is reported*		

Discussion					
Key results	18	Summarise key results with reference to study objectives	*Covered in discussion*		
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	*An extensive limitations section is included, covering the relevant aspects*	RECORD 19.1: Discuss the implications of using data that were not created or collected to answer the specific research question(s). Include discussion of misclassification bias, unmeasured confounding, missing data, and changing eligibility over time, as they pertain to the study being reported.	*An extensive limitations section is included, covering the relevant aspects*
Interpretation	20	Give a cautious overall interpretation of results considering objectives,	*Covered in discussion*		
		limitations, multiplicity of analyses, results from similar studies, and other relevant evidence			
Generalisability	21	Discuss the generalisability (external validity) of the study results	*Covered in discussion*		
Other Information					
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present	*Covered in funding section*		

		article is based			
Accessibility of protocol, raw data, and programming code		..		RECORD 22.1: Authors should provide information on how to access any supplemental information such as the study protocol, raw data, or programming code.	*Information is included in the data availability statement*

*Reference: Benchimol EI, Smeeth L, Guttmann A, Harron K, Moher D, Petersen I, Sørensen HT, von Elm E, Langan SM, the RECORD Working Committee. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. *PLoS Medicine* 2015; in press.

BMJ Open

Severity of COVID-19 and adverse long-term outcomes: a retrospective cohort study based on a US electronic health record database

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-056284.R2
Article Type:	Original research
Date Submitted by the Author:	16-Nov-2021
Complete List of Authors:	Jovanoski, Nick; F Hoffmann-La Roche Ltd Chen, Xin; Genentech Inc Becker, Ursula; F Hoffmann-La Roche Ltd Zalocusky, Kelly; Genentech Inc Chawla, Devika; Genentech Inc Tsai, Larry; Genentech Inc Borm, Michelle; F Hoffmann-La Roche Ltd Neighbors, Margaret; Genentech Inc Yau, Vincent; Genentech Inc
Primary Subject Heading:	Respiratory medicine
Secondary Subject Heading:	Epidemiology
Keywords:	COVID-19, Epidemiology < INFECTIOUS DISEASES, PUBLIC HEALTH, Respiratory infections < THORACIC MEDICINE

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Severity of COVID-19 and adverse long-term outcomes: a retrospective cohort study based on a US electronic health record database

Nick Jovanoski[1], Xin Chen[2], Ursula Becker[1], Kelly Zalocusky[2], Devika Chawla[2], Larry Tsai[2], Michelle Borm[3], Margaret Neighbors[2], Vincent Yau[2]

1. F. Hoffmann-La Roche Ltd, Basel, Switzerland

2. Genentech, Inc., South San Francisco, CA, USA

3. Roche Nederland BV, Woerden, Netherlands

Author for correspondence: Vincent Yau, Ph.D. Genentech, Inc., South San Francisco, CA, USA. yau.vincent@gene.com

Target journal: BMJ Open

Word count: 3726/4000

Figures / Tables: 4 / 2 (plus 1 / 6 supplementary)

References: 29

Abstract (300/300 words)

Objective: To identify potential risk factors for adverse long-term outcomes (LTOs) associated with COVID-19, using a large electronic health record (EHR) database.

Design: Retrospective cohort study. Patients with COVID-19 were assigned into sub-cohorts according to most intensive treatment setting experienced. Newly diagnosed conditions were classified as respiratory, cardiovascular, or mental health LTOs at >30–≤90 or >90–≤180 days after COVID-19 diagnosis or hospital discharge. Multivariate regression analysis was performed to identify any association of treatment setting (as a proxy for disease severity) with LTO incidence.

Setting: Optum® de-identified COVID-19 EHR dataset drawn from hospitals and clinics across the United States.

Participants: Individuals diagnosed with COVID-19 (N=57,748) from February 20–July 4, 2020.

Main outcomes: Incidence of new clinical conditions after COVID-19 diagnosis or hospital discharge and the association of treatment setting (as a proxy for disease severity) with their risk of occurrence.

Results: Patients were assigned into one of six sub-cohorts: outpatient (n=22,788), emergency room (ER) with same-day COVID-19 diagnosis (n=11,633), ER with COVID-19 diagnosis ≤21 days before ER visit (n=2,877), hospitalization without intensive care unit (ICU; n=16,653), ICU without ventilation (n=1,837), and ICU with ventilation (n=1,960). Respiratory LTOs were more common than cardiovascular or mental health LTOs across sub-cohorts, and LTO incidence was higher in hospitalized versus non-hospitalized sub-cohorts. Patients with the most severe disease were at increased risk of respiratory (risk ratio [RR] 1.86, 95% confidence interval [CI] 1.56, 2.21), cardiovascular (RR 2.65, 95% CI 1.49, 4.43), and mental health outcomes (RR 1.52, 95% CI 1.20, 1.91) up to six months after hospital discharge compared with outpatients.

Conclusions: Patients with severe COVID-19 had increased risk of new clinical conditions up to six months after hospital discharge. The extent that treatment setting (e.g., ICU) contributed to these conditions is unknown, but strategies to prevent COVID-19 progression may nonetheless minimize their occurrence.

Strengths and limitations of this study

- This study used a large electronic health record database containing a rich source of patient-level medical and administrative records from hospitals, emergency departments, and outpatient centers across the United States.
- Multivariate logistic regression analysis was used to adjust for measured confounders and assess the association of increasing COVID-19 severity (proxied by treatment setting) with the risk of new clinical conditions being diagnosed up to six months after COVID-19 diagnosis or hospital discharge.
- A sensitivity analysis assessing the association of increasing COVID-19 severity (proxied by treatment setting) with the risk of a new cancer diagnosis served as a negative control.
- The main limitation of this retrospective study is that we use treatment setting as a proxy for COVID-19 severity, and therefore it is difficult to tease out associations specific to the treatment setting (e.g., invasive ventilation) from the underlying COVID-19 severity; any differences that exist between cohorts could bias the results, and as all potential confounders may not be controlled for, the results do not indicate causality.
- Additional limitations include missing information on smoking status, the lack of a COVID-19-negative control group, the possibility of missing data, being restricted to examining conditions captured by ICD-10 codes, the lack of information on COVID-

19 treatments received, and the lack of laboratory values or other biomarkers to better characterize disease.

Introduction

The coronavirus disease 2019 (COVID-19) pandemic caused by the novel severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) has imposed an immense burden of morbidity and mortality worldwide.[1] Although the majority of patients experience mild or moderate symptoms that resolve within a few weeks of initial infection, increasing evidence suggests that a subset of patients continue to display symptoms beyond four weeks after infection.[2 3] These symptoms are wide ranging and often extend beyond the typical initial symptoms of COVID-19 to include respiratory (e.g., dyspnea, decreased exercise capacity), cardiovascular (e.g., heart palpitations, chest pain), and mental health (e.g., confusion, disorientation) disorders.[4 5] Notably, such outcomes have been observed even in patients with mild acute COVID-19 symptoms.[6] These prolonged symptoms have collectively been referred to by several names including post-acute COVID-19 (PAC), post-COVID-19 syndrome (PCS), post-acute sequelae of SARS-CoV-2 infection (PASC), and possibly more commonly 'long COVID'.[7 8] However, due to the overlapping and non-specific range of symptoms experienced, the medical community has not yet converged on precise definitions, and it is possible that distinct subsets of long COVID patients exist. It has also been suggested that long COVID can be further sub-divided into subacute COVID-19 (4–12 weeks after initial onset of COVID-19 symptoms) and post-COVID-19 syndrome (beyond 12 weeks).[4 9] The underlying pathogenic mechanisms of long COVID are not well understood, but multiple causes have been proposed, including immune dysregulation and viral persistence.[10] Additionally, in patients with severe disease requiring treatment in the intensive care unit (ICU), non-specific secondary effects cannot be ruled out, similar to those observed in 'post-intensive care syndrome'.[11]

High-quality clinical data on respiratory, cardiovascular, and neurologic sequelae of SARS-CoV-2 infection are beginning to emerge,[12-14] and several observational studies and patient registries have been established to better understand the long-term outcomes (LTOs)

of COVID-19.[15 16] However, little is known about the potential baseline factors that may predict the development of long COVID.

Retrospective cohort studies using electronic health records (EHRs) are uniquely positioned due to their size and convenience to provide insights into factors underlying long COVID development and the range of long COVID conditions that exist. The Optum® de-identified COVID-19 EHR dataset contains patient-level medical and administrative records from hospitals, emergency departments, outpatient centers, and laboratories across the United States (US). This dataset has previously been utilized to describe key epidemiological features of a large cohort of hospitalized patients with COVID-19[17] and to develop a prognostic model of in-hospital mortality.[18]

The current study utilized the Optum® de-identified COVID-19 EHR dataset to better understand the types of LTOs encountered by patients with long COVID, to define the factors that predict their diagnosis, and to understand the role that treatment setting (as a proxy for COVID-19 severity) plays in the manifestation of these outcomes.

Methods

Database

Individuals with COVID-19 diagnosed between February 20, 2020 and July 4, 2020 were extracted from the Optum® de-identified COVID-19 EHR dataset (569,149 individuals from 3,832,315 in the entire dataset). This dataset contains patient-level medical and administrative records from hospitals, emergency departments, outpatient centers, and laboratories across the US. All data were de-identified according to the Health Insurance Portability and Accountability Act Expert Method and managed according to Optum® customer data use agreements. The COVID-19 EHR dataset comprises clinical information sourced from hospital networks that provided data meeting Optum®'s internal data quality criteria. Data cleaning methods used were as described previously.[17]

Patients and study design

Eligible patients (overall COVID-19 cohort) had ≥1 of the following: a COVID-19 diagnosis code (U07.1, U07.2), a positive diagnostic test for SARS-CoV-2 infection (e.g., molecular or antigen test), or a B97.29 diagnosis code (other coronavirus as the cause of diseases classified elsewhere) without a negative SARS-CoV-2 molecular test within 14 days. The index date was defined as the date of COVID-19 diagnosis or COVID-19-related hospitalization (as defined below), whichever occurred first. The baseline period was defined as the 12 months prior to the index date, and a minimum of 180 days follow-up was required for all patients. The overall study design is shown in **Figure 1**.

Eligible patients were assigned into the following six sub-cohorts according to treatment setting: **1. Outpatient**, patients with a COVID-19 diagnosis and no record of hospitalization or an emergency room (ER) visit within 21 days of diagnosis; **2. ER on diagnosis**, COVID-19 diagnosis on the same day as ER visit; **3. ER**, COVID-19 diagnosis prior to ER visit, i.e., patients with an ER visit within 21 days after COVID-19 diagnosis (excluding diagnosis date); **4. Hospitalization without ICU**, patients hospitalized with no

record of ICU admission; **5. Hospitalized with ICU but no ventilation**, patients hospitalized with record of ICU admission but no record of ventilator or extracorporeal membrane oxygenation (ECMO) use during ICU stay; **6. Hospitalized with ICU and ventilation**, patients hospitalized with record of ICU admission and ventilator or ECMO use during ICU stay.

Hospitalization was defined as an inpatient or ER overnight visit with an initial COVID-19 diagnosis made during hospitalization and within seven days of admission, or an inpatient or ER overnight visit within 21 days of the initial COVID-19 diagnosis, where the hospital had a record of this diagnosis. Contiguous ER and inpatient visits with a gap of up to one day were considered a single hospitalization. If a patient had multiple eligible hospitalizations, only data from the first hospitalization were considered, as described previously.[17]

Modeling and statistical analysis

LTOs occurring >30–≤180 days after hospital discharge or COVID-19 diagnosis were categorized into one of two time windows (>30–≤90 days or >90–≤180 days) and were further classified as respiratory, cardiovascular, or mental health conditions (**Supplemental Table 1**).[19] LTOs were selected to capture a broad range of potential sequelae, even if there was no strong clinical or pathological rationale for their choice, given the absence of sufficient clinical data regarding established complications associated with COVID-19. Multivariate logistic regression analyses were performed to determine the association of disease severity (proxied by treatment setting) with the three LTO classifications. Covariates were intended to encompass the main known risk factors for developing severe COVID-19,[20] and included demographic information (i.e., age, gender, race, ethnicity, diagnosis month, insurance type, obesity status) and baseline health conditions (i.e., those included in the Charlson Comorbidity Index [CCI] (**Supplemental Table 2**). CCI was treated as a numeric variable, while all other variables were treated as categorical. Age was binned into <18 years, 18–29, 30–39, 40–49, 50–65, 65–74, 75–84, and ≥85 years. Date of diagnosis was

also binned into months in 2020 (pre-April, April, May, June, July; allowing for ≥180 days follow-up until 31 December 2020 at the latest). Patients were excluded from the regression model examining a specific LTO category if they had a diagnosis in that category in the 12 months prior to the index date (for example, if a patient had an asthma diagnosis 12 months prior to the index date, they would be excluded from the model for respiratory LTOs).

All statistical analysis was performed using R 3.6.3.[21] Using the sjstats package, regression was performed using the function 'glm' and the risk ratio (RR) was calculated by converting the odds ratio (OR) using the function 'OR to RR'.[22] Increased risk of diagnosis of a health condition was implied when the RR and both the low and high 95% confidence interval limits (CI) were >1, and decreased risk was implied when the RR and low and high 95% CIs were <1.

Sensitivity analysis

A sensitivity analysis was performed to investigate the potential association of disease severity (proxied by treatment setting) with risk of a new cancer diagnosis, to serve as a negative control. The same set of covariates was used as per the main analysis, but cancer diagnosis was the only LTO examined. Currently, no evidence exists to suggest that COVID-19 severity increases the risk of a new cancer diagnosis. Thus, an association here may indicate that the associations from the main analysis may be driven by other differences between patients across treatment settings.

Patient and Public Involvement

No patient involved.

Results

Patient population

In total, 57,748 patients were eligible for the overall COVID-19 cohort. **Table 1** presents descriptive statistics of the patients by sub-cohort. Mean age tended to be higher in patients in hospitalized sub-cohorts (53.2–57.7 years) than in those in non-hospitalized sub-cohorts (41.0–46.8 years). Overall, 53.3% of patients were female. Across all patients, 50.3% were Caucasian, 22.8% were African American, 3.2% were Asian, and the remaining 23.6% were missing information on race. Additionally, 67.5% were of non-Hispanic ethnicity, while data on ethnicity was missing for 11.8% of patients. Overall, 19% of patients were obese and the mean weighted CCI score was 1.20. Information on smoking status was missing for 93.1% of patients (**Table 1**). Full details of demographics and baseline characteristics are provided in **Supplemental Table 3**.

The proportions of patients with incipient respiratory, cardiovascular, and / or mental health conditions that were diagnosed either >30–≤90 days or >90–≤180 days after COVID-19 diagnosis or hospital discharge are provided in **Table 2**. The proportions of patients with new LTOs were generally higher in the sub-cohorts with more severe disease (i.e., the ER sub-cohort and all hospitalized sub-cohorts) compared with the outpatient sub-cohort. In addition, the proportion of patients with respiratory LTOs was higher than the proportions with cardiovascular or mental health LTOs. New respiratory LTOs were diagnosed more frequently during the earlier time window across sub-cohorts, except in the outpatient sub-cohort where the proportion of patients diagnosed was the same in both time windows (both 8.1%; **Table 2**). No clear temporal trends were noted for diagnosis of cardiovascular or mental health LTOs, with similar proportions of patients with new cardiovascular and mental health LTOs observed in the >30–≤90- and >90–≤180-day windows for each sub-cohort (**Table 2**). The proportions of patients with LTOs in more than one category (i.e., 'respiratory and cardiovascular', 'respiratory and mental health', 'mental health and cardiovascular', or 'respiratory, cardiovascular, and mental health') were lower than the proportions of patients

with LTOs in a single category, suggesting that a diagnosis in one category did not necessarily lead to a diagnosis in another.

Regarding individual conditions, the prevalence of newly diagnosed pneumonia, dyspnea, and respiratory failure in the >90–≤180-day window closely followed the pattern of initial COVID-19 severity (as proxied by treatment setting), with most cases being diagnosed in the 'ICU with ventilation' sub-cohort (**Supplemental Table 4**). Similarly, although encephalopathy, confusion or disorientation, cardiac arrhythmia, and myocardial infarction were less common, the prevalence of these conditions also increased with increasing COVID-19 severity. Full details of conditions that were diagnosed in the >30–≤90- and >90–≤180-day windows following COVID-19 diagnosis or hospital discharge are provided in **Supplemental Table 4**.

Modeling

The most striking potential covariate associated with increased risk of newly diagnosed respiratory conditions at >30–≤90 days and >90–≤180 days post COVID-19 diagnosis or hospital discharge was increasing severity of illness according to increasing hospitalization severity, utilizing the outpatient sub-cohort as the reference group (**Figure 2** and **Supplemental Table 5**). ICU with ventilation was associated with increased risk of a novel respiratory condition diagnosis compared with the outpatient sub-cohort at >30–≤90 days (RR 2.64, 95% CI 2.27, 3.04) and >90–≤180 days post COVID-19 diagnosis or hospital discharge (RR 1.86, 95% CI 1.55, 2.21); in addition, ICU without ventilation was associated with increased risk during the >30–≤90-day time window (RR 1.69, 95% CI 1.39, 2.03), while ER was associated with increased risk at both >30–≤90 days (RR 1.39, 95% CI 1.17, 1.65) and >90–≤180 days (RR 1.33, 95% CI 1.10, 1.58) post COVID-19 diagnosis or hospital discharge. By contrast, patients with an ER visit on the COVID-19 diagnosis date were less likely than those in the outpatient sub-cohort to be diagnosed with a new respiratory condition at >30–≤90 days (RR 0.64, 95% CI 0.56, 0.74) and 90–180 days post COVID-19 diagnosis or hospital discharge (RR 0.56, 95% CI 0.48, 0.65). Additional covariates

associated with increased risk of new respiratory conditions were older patient age and obesity. A COVID-19 diagnosis during or prior to April 2020 exhibited a non-significant trend towards increased risk of new respiratory condition occurrence compared with later diagnosis, which may reflect changes in treatment algorithms over time. Full results are presented in **Supplemental Table 5**.

Increasing hospitalization severity was also found to be associated with increased risk of a new cardiovascular condition occurring post COVID-19 diagnosis or hospital discharge (**Figure 3** and **Supplemental Table 5**). Notably, ICU with ventilation was associated with increased risk of the occurrence of novel cardiovascular conditions compared with the outpatient sub-cohort at >30–≤90 days (RR 3.16, 95% CI 1.83, 5.18) and >90–≤180 days post COVID-19 diagnosis or hospital discharge (RR 2.65, 95% CI 1.49, 4.43), while ICU without ventilation was associated with increased risk during the >90–≤180-day time window (RR 2.41, 95% CI 1.25, 4.23). Similar to the findings regarding respiratory conditions, patients with an ER visit on the COVID-19 diagnosis date were less likely than outpatients to be diagnosed with novel cardiovascular conditions in both the >30–≤90-day (RR 0.45, 95% CI 0.27, 0.71) and >90–≤180-day windows (RR 0.59, 95% CI 0.38, 0.89). Additional covariates associated with an increased risk of new cardiovascular conditions occurring included older patient age and non-Hispanic ethnicity. Full results are presented in **Supplemental Table 5**.

The risk of a new mental health condition occurring post COVID-19 diagnosis or hospital discharge also increased according to increasing hospitalization severity (**Figure 4** and **Supplemental Table 5**). ICU with ventilation was associated with increased risk of a new mental health condition occurring compared with the outpatient sub-cohort at >30–≤90 days (RR 1.89, 95% CI 1.51, 2.35) and >90–≤180 days post COVID-19 diagnosis or hospital discharge (RR 1.52, 95% CI 1.20, 1.91), and ICU without ventilation was similarly associated with increased risk of a new mental health condition diagnosis during the >90–≤180-day window (RR 1.34, 95% CI 1.02, 1.73). Of note, compared with those <18 years, all age groups examined appeared to be at higher risk of the occurrence of new mental health

conditions at >30–≤90 days post COVID-19 diagnosis or hospital discharge. In the >90–≤180-day window, only the 65–74 and 75–84 years age groups were not at higher risk. Additional covariates associated with increased risk of a new mental health condition occurring included obesity, Caucasian race, and non-Hispanic ethnicity. See **Supplemental Table 5** for full results.

Sensitivity analysis

With the exception of older age, COVID-19 severity (proxied by treatment setting) did not predict a new cancer diagnosis up to 180 days after COVID-19 diagnosis or hospital discharge (**Supplemental Figure 1** and **Supplemental Table 6**), giving confidence in the results of the original analysis.

Discussion

By utilizing EHRs of over 55,000 patients from hospitals and clinics across the US, this study set out to examine the types of new LTOs (i.e., only those that were identified after COVID-19 diagnosis or hospital discharge) associated with long COVID and to identify potential underlying factors that may contribute to their occurrence. Severe disease was found to predict an increased likelihood of a new LTO diagnosis, whereby increasing hospitalization severity was associated with increased risk of new respiratory (e.g., pneumonia), cardiovascular (e.g., myocardial infarction), and mental health conditions (e.g., confusion or disorientation). In severely affected COVID-19 patients, some LTOs were diagnosed between three and six months after hospital discharge, suggesting that the overall COVID-19 burden extends far beyond the acute infection phase. In addition, although patients with severe disease were most at risk of presenting with new LTOs, non-hospitalized patients also experienced a relatively high incidence of LTOs, suggesting that even patients with mild disease are at risk of adverse long-term effects associated with COVID-19.

Although the data show a clear general trend of increased LTOs that correlated with COVID-19 severity (proxied by treatment setting), the specificity of this effect to COVID-19 is unclear, as ICU survivors commonly develop a range of new conditions upon discharge collectively referred to as 'post-intensive care syndrome', regardless of their underlying diagnosis.[11] Nonetheless, preventing the development of more severe disease, where possible, may decrease the likelihood of health problems post infection and would be expected to simultaneously increase the probability of survival. Together, these effects would have a cumulative positive impact on both patients and healthcare systems.

Interestingly, the 'ER on diagnosis' sub-cohort exhibited a reduced incidence of LTOs compared with the outpatient sub-cohort. The reasons for this are not clear but are likely due in part to the lower mean age and reduced incidence of comorbidities in this sub-cohort relative to the other sub-cohorts. In addition, it is possible that in the context of the pandemic, when primary care physicians had more limited personal protective equipment

and other resources, these patients were directed to the ER to be tested for COVID-19, despite not having severe enough disease to warrant an ER visit. Finally, depending on the hospital setting and processes in place, asymptomatic patients who attended the ER for non-COVID-19 reasons may have tested positive while there, which may have led to the inclusion of milder COVID-19 cases in this sub-cohort.

Previous studies have examined the link between COVID-19 severity and LTOs. A study of 2,469 hospitalized COVID-19 patients in Wuhan, China showed that more severe disease correlated with increased risk of LTOs up to six months after infection, including fatigue, sleep difficulties, and anxiety or depression.[23] Anxiety or depression was observed in 23% of patients in that study compared with ~10% in our study; this difference is likely because our study was limited to newly diagnosed disorders in both inpatients and outpatients, while the previous study included new or worsening symptoms in hospitalized patients only. A separate, large study of COVID-19 patients that utilized a US EHR database (N=236,379) to examine six-month outcomes (inpatients and outpatients) reported that ~7% of patients had a first anxiety disorder compared with ~17% that had any anxiety disorder, and that increased incidence was correlated with increased disease severity.[13] A further study compared 73,435 non-hospitalized COVID-19 patients who were users of the Veterans Health Administration with 4,990,835 control patients and reported an increased risk of incident sequelae including, but not limited to, respiratory, cardiovascular, and mental health disorders after a median follow-up duration of 126 and 130 days, respectively.[19] Smaller, single-site hospital studies in the United Kingdom have reported similar trends between disease severity and shorter-term outcomes, with breathlessness commonly reported up to 12 weeks post COVID-19.[24 25] In addition, self-reported data in patients with COVID-19 (N=4,182) showed that upper respiratory complaints (e.g., shortness of breath) and cardiac symptoms (e.g., palpitations, tachycardia) were commonly reported in patients with long COVID (symptoms lasting ≥28 days),[26] and data from a separate study utilizing wearable devices provided further evidence of prolonged tachycardia in symptomatic patients with COVID-19.[27] The current study builds on these previous reports and provides additional

evidence of a link between COVID-19 severity (proxied by treatment setting) and increased risk of developing LTOs, using a large dataset from both hospitalized and non-hospitalized patients. In addition, our study provides a detailed summary of the incidence of a wide range of specific health conditions that occurred up to six months after COVID-19 diagnosis or hospital discharge, providing a useful resource to better understand and characterise the range of conditions that constitute long COVID.

Our study categorized three major classes of LTOs that occur in patients with long COVID: respiratory, cardiovascular, and mental health. This is broadly in keeping with a previous retrospective cohort study in England that followed 48,780 patients hospitalized with COVID-19, who had significantly higher rates of respiratory and cardiovascular disease after a mean follow-up of 140 days.[28] In addition, a retrospective study that used a large administrative all-payor database including 27,589 inpatients and 46,857 outpatients demonstrated that post COVID-19, patients were more likely to experience a range of conditions, including respiratory, nervous, and circulatory system conditions, than outpatient control patients.[29] A greater understanding of the conditions that characterize long COVID is needed to better anticipate the future healthcare burden of COVID-19 and to optimize strategies to minimize long COVID development. In this regard, signals detected in the current study such as lung fibrosis, as well as other factors including pediatric long COVID, vaccination effects, and healthcare utilization, are topics that may warrant future analysis. In particular, a greater understanding of the long-term economic consequences of COVID-19 and the impact of long COVID on patient quality of life is needed.

A major limitation of this analysis is that treatment setting is used as a proxy for COVID-19 severity; therefore, it is difficult to tease out the effect of treatment setting procedures (e.g., invasive ventilation) from the underlying COVID-19 severity. Furthermore, our analysis did not distinguish short-term from chronic health conditions. Additional limitations include missing information on smoking status, the restriction of follow-up to only six months, the lack of a COVID-19-negative control group, the possibility of missing data

(e.g., patients may have sought care for an LTO not captured in the Optum® de-identified COVID-19 EHR dataset), the lack of information on COVID-19 treatments received, and the lack of laboratory values or other biomarkers to better characterize disease. Finally, capture of health conditions relies on International Classification of Disease-10 (ICD-10) codes, whereas some conditions of interest (e.g., anosmia, ageusia, and brain fog) lack specific ICD-10 codes and other conditions are known to be under-captured. The B97.29 diagnosis code includes other coronaviruses in addition to SARS-CoV-2 and may therefore be a potential limitation of our study; however, the majority of our COVID-19 cohort (>85%) was diagnosed from April to July using the official U07.1 diagnosis code that is specific to COVID-19, meaning it is unlikely that a substantial number of infections, if any, were from other coronaviruses.

Conclusions

Although LTOs were reported in patients across all sub-cohorts, increased risk of new respiratory, cardiovascular, and mental health conditions was observed with increasing COVID-19 severity, using treatment setting as a proxy. Strikingly, the risk of new conditions being diagnosed remained high up to six months post COVID-19 diagnosis or hospital discharge, suggesting that the burden of COVID-19 extends far beyond the acute infection phase. Future research is warranted to understand specific factors that lead to the occurrence of new LTOs in patients with COVID-19, and to distinguish between the relative effect of COVID-19 severity versus any general effects that may occur after acute critical illness.

Acknowledgements

We thank Shemra Rizzo for valuable contributions. Third-party medical writing assistance, under the direction of the authors, was provided by John Bett, PhD, of Ashfield MedComms, an Ashfield Health company, and funded by F. Hoffmann-La Roche Ltd.

Contributors

All authors were involved in drafting and revising the manuscript, approved the final version, and agree to being accountable for all aspects of the work. Nick Jovanoski contributed to the conception of the research question, study design, analysis, and data interpretation. Xin Chen contributed to study design, analysis and data interpretation. Ursula Becker contributed to the conception of the research question, study design, analysis, and data interpretation. Kelly Zalocusky contributed to the conception of the research question, design of the analysis, selection of outcomes and data interpretation. Devika Chawla contributed to the conception of the research question, design of the analysis, and selection of outcomes. Larry Tsai contributed to study design and data interpretation. Michelle Borm contributed to data interpretation. Margaret Neighbors contributed to selection and categorization of key complications for study design. Vincent Yau contributed to the study design, acquisition, analysis and data interpretation.

Funding statement

This work was supported by F. Hoffmann-La Roche Ltd.

Competing interests statement

Nick Jovanoski and Ursula Becker are employees of F. Hoffmann-La Roche Ltd. Michelle Borm is an employee of Roche Nederland BV. Ursula Becker and Michelle Borm hold shares

in F. Hoffmann-La Roche Ltd. Xin Chen, Devika Chawla, Larry Tsai, Margaret Neighbors, and Vincent Yau are employees of Genentech, Inc. and hold shares in F. Hoffmann-La Roche Ltd. Kelly Zalocusky is a former employee of Genentech, Inc. and holds shares in F. Hoffmann-La Roche Ltd.

Patient consent

None required

Ethics approval

The use of the Optum® de-identified COVID-19 EHR dataset was reviewed by the New England Institutional Review Board (IRB) and was determined to be exempt from broad IRB approval, as this study did not involve human subject research.

Data availability statement

Data may be obtained from a third party and are not publicly available. Data were licensed from Optum® and interested researchers may contact Optum® for data access requests. All interested researchers can access the data in the same manner as the authors. The authors had no special access privileges.

References

1. Centers for Disease Control and Prevention. Estimated Disease Burden of COVID-19 2021 [Available from: https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/burden.html accessed May 2021.
2. Carfi A, Bernabei R, Landi F, et al. Persistent Symptoms in Patients After Acute COVID-19. *JAMA* 2020;324(6):603-05. doi: 10.1001/jama.2020.12603 [published Online First: 2020/07/10]
3. Tenforde MW, Kim SS, Lindsell CJ, et al. Symptom Duration and Risk Factors for Delayed Return to Usual Health Among Outpatients with COVID-19 in a Multistate Health Care Systems Network - United States, March-June 2020. *MMWR Morb Mortal Wkly Rep* 2020;69(30):993-98. doi: 10.15585/mmwr.mm6930e1 [published Online First: 2020/07/31]
4. Nalbandian A, Sehgal K, Gupta A, et al. Post-acute COVID-19 syndrome. *Nat Med* 2021;27(4):601-15. doi: 10.1038/s41591-021-01283-z [published Online First: 2021/03/24]
5. Centers for Disease Control and Prevention. Post-COVID Conditions 2021 [Available from: https://www.cdc.gov/coronavirus/2019-ncov/long-term-effects.html accessed May 2021.
6. Augustin M, Schommers P, Stecher M, et al. Post-COVID syndrome in non-hospitalised patients with COVID-19: a longitudinal prospective cohort study. *Lancet Reg Health Eur* 2021;6:100122. doi: 10.1016/j.lanepe.2021.100122 [published Online First: 2021/05/25]
7. Mullard A. Long COVID's long R&D agenda. *Nat Rev Drug Discov* 2021;20(5):329-31. doi: 10.1038/d41573-021-00069-9 [published Online First: 2021/04/22]
8. National Institutes of Health. NIH launches new initiative to study “Long COVID” Bethesda, Maryland: National Institutes of Health; 2021 [Available from: https://www.nih.gov/about-nih/who-we-are/nih-director/statements/nih-launches-new-initiative-study-long-covid accessed July 2021.

9. Mayor N, Tsang R, Joy M, et al. Long covid: coding is caring. *BMJ* 2021;373:n1262. doi: 10.1136/bmj.n1262 [published Online First: 2021/05/21]

10. Marx V. Scientists set out to connect the dots on long COVID. *Nat Methods* 2021;18(5):449-53. doi: 10.1038/s41592-021-01145-z [published Online First: 2021/04/30]

11. Brown SM, Bose S, Banner-Goodspeed V, et al. Approaches to Addressing Post-Intensive Care Syndrome among Intensive Care Unit Survivors. A Narrative Review. *Ann Am Thorac Soc* 2019;16(8):947-56. doi: 10.1513/AnnalsATS.201812-913FR [published Online First: 2019/06/05]

12. Myall KJ, Mukherjee B, Castanheira AM, et al. Persistent Post-COVID-19 Interstitial Lung Disease. An Observational Study of Corticosteroid Treatment. *Ann Am Thorac Soc* 2021;18(5):799-806. doi: 10.1513/AnnalsATS.202008-1002OC [published Online First: 2021/01/13]

13. Taquet M, Geddes JR, Husain M, et al. 6-month neurological and psychiatric outcomes in 236 379 survivors of COVID-19: a retrospective cohort study using electronic health records. *Lancet Psychiatry* 2021;8(5):416-27. doi: 10.1016/S2215-0366(21)00084-5 [published Online First: 2021/04/10]

14. Puntmann VO, Carerj ML, Wieters I, et al. Outcomes of Cardiovascular Magnetic Resonance Imaging in Patients Recently Recovered From Coronavirus Disease 2019 (COVID-19). *JAMA Cardiol* 2020;5(11):1265-73. doi: 10.1001/jamacardio.2020.3557 [published Online First: 2020/07/31]

15. University of Birmingham. Why are some people more susceptible to the long-term health effects of COVID-19? 2021 [Available from: https://www.birmingham.ac.uk/research/quest/21st-century-healthcare/long-covid.aspx accessed May 2021.

16. Mount Sinai Hospital. Center for Post-COVID Care 2021 [Available from: https://www.mountsinai.org/about/covid19/center-post-covid-care accessed May 2021.

17. Chawla D, Rizzo S, Zalocusky K, et al. Descriptive epidemiology of 16,780 hospitalized COVID-19 patients in the United States. *medRxiv* 2021 doi: https://doi.org/10.1101/2020.07.17.20156265

18. Incerti D, Rizzo S, Li X, et al. Prognostic model to identify and quantify risk factors for mortality among hospitalised patients with COVID-19 in the USA. *BMJ open* 2021;11(4):e047121.

19. Al-Aly Z, Xie Y, Bowe B. High-dimensional characterization of post-acute sequelae of COVID-19. *Nature* 2021 doi: 10.1038/s41586-021-03553-9

20. Centers for Disease Control and Prevention. People with Certain Medical Conditions 2021 [Available from: https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/people-with-medical-conditions.html accessed July 2021.

21. R Core Team. R: A language and environment for statistical computing Vienna, Austria2020 [Available from: https://www.r-project.org/.

22. Lüdecke D. Statistical Functions for Regression Models (Version 0.18.1) 2021 [Available from: https://CRAN.R-project.org/package=sjstats.

23. Huang C, Huang L, Wang Y, et al. 6-month consequences of COVID-19 in patients discharged from hospital: a cohort study. *Lancet* 2021;397(10270):220-32. doi: 10.1016/S0140-6736(20)32656-8 [published Online First: 2021/01/12]

24. Halpin SJ, McIvor C, Whyatt G, et al. Postdischarge symptoms and rehabilitation needs in survivors of COVID-19 infection: A cross-sectional evaluation. *J Med Virol* 2021;93(2):1013-22. doi: 10.1002/jmv.26368 [published Online First: 2020/07/31]

25. Arnold DT, Hamilton FW, Milne A, et al. Patient outcomes after hospitalisation with COVID-19 and implications for follow-up: results from a prospective UK cohort. *Thorax* 2020 doi: 10.1136/thoraxjnl-2020-216086 [published Online First: 2020/12/05]

26. Sudre CH, Murray B, Varsavsky T, et al. Attributes and predictors of long COVID. *Nature Medicine* 2021;27(4):626-31. doi: 10.1038/s41591-021-01292-y

27. Radin JM, Quer G, Ramos E, et al. Assessment of Prolonged Physiological and Behavioral Changes Associated With COVID-19 Infection. *JAMA Netw Open*

2021;4(7):e2115959. doi: 10.1001/jamanetworkopen.2021.15959 [published Online First: 2021/07/08]

28. Ayoubkhani D, Khunti K, Nafilyan V, et al. Post-covid syndrome in individuals admitted to hospital with covid-19: retrospective cohort study. *BMJ* 2021;372:n693. doi: 10.1136/bmj.n693 [published Online First: 2021/04/02]

29. Chevinsky JR, Tao G, Lavery AM, et al. Late conditions diagnosed 1-4 months following an initial COVID-19 encounter: a matched cohort study using inpatient and outpatient administrative data - United States, March 1-June 30, 2020. *Clin Infect Dis* 2021 doi: 10.1093/cid/ciab338 [published Online First: 2021/04/29]

Tables and Figures

Table 1 Baseline characteristics of COVID-19 patients overall and by sub-cohort

		Sub-cohort					
	All patients (N=57,748)	**1. Outpatient (n=22,788)**	**2. ER on diagnosis (n=11,633)**	**3. ER (n=2,877)**	**4. Hospitalization without ICU (n=16,653)**	**5. ICU without ventilation (n=1,837)**	**6. ICU with ventilation (n=1,960)**
Mean age (SD), years	47.93 (18.76)	46.78 (18.90)	40.95 (16.89)	44.00 (16.61)	53.17 (18.50)	55.92 (17.32)	57.70 (14.78)
Age group, n (%)							
<18 years	2,184 (3.8)	1,033 (4.5)	641 (5.5)	78 (2.7)	366 (2.2)	46 (2.5)	20 (1.0)
18–29 years	8,509 (14.7)	3,693 (16.2)	2,543 (21.9)	538 (18.7)	1,574 (9.5)	89 (4.8)	72 (3.7)
30–39 years	8,972 (15.5)	3,541 (15.5)	2,496 (21.5)	579 (20.1)	2,046 (12.3)	175 (9.5)	135 (6.9)
40–49 years	9,362 (16.2)	3,664 (16.1)	2,205 (19.0)	555 (19.3)	2,442 (14.7)	253 (13.8)	243 (12.4)
50–64 years	16,103 (27.9)	6,231 (27.3)	2,706 (23.3)	780 (27.1)	4,937 (29.6)	626 (34.1)	823 (42.0)
65–74 years	7,065 (12.2)	2,658 (11.7)	689 (5.9)	230 (8.0)	2,683 (16.1)	363 (19.8)	442 (22.6)
75–84 years	3,620 (6.3)	1,279 (5.6)	239 (2.1)	76 (2.6)	1,647 (9.9)	195 (10.6)	184 (9.4)
≥85 years	891 (1.5)	303 (1.3)	54 (0.5)	22 (0.8)	440 (2.6)	44 (2.4)	28 (1.4)
Missing	1,042 (1.8)	386 (1.7)	60 (0.5)	19 (0.7)	518 (3.1)	46 (2.5)	13 (0.7)
Sex, n (%)							

Female	30,782 (53.3)	12,856 (56.4)	6,115 (52.6)	1,721 (59.8)	8,487 (51.0)	829 (45.1)	774 (39.5)
Male	26,939 (46.6)	9,920 (43.5)	5,515 (47.4)	1,152 (40.0)	8,160 (49.0)	1,008 (54.9)	1,184 (60.4)
Missing	27 (<0.1)	12 (<0.1)	3 (<0.1)	4 (<0.1)	6 (<0.1)	0 (0.0)	2 (<0.1)
Race, n (%)							
African American	13,183 (22.8)	3,473 (15.2)	3,178 (27.3)	790 (27.5)	4,675 (28.1)	551 (30.0)	516 (26.3)
Asian	1,848 (3.2)	639 (2.8)	438 (3.8)	100 (3.5)	555 (3.3)	41 (2.2)	75 (3.8)
Caucasian	29,074 (50.3)	13,746 (60.3)	4,653 (40.0)	1,337 (46.5)	7,538 (45.3)	849 (46.2)	951 (48.5)
Missing	13,643 (23.6)	4,930 (21.6)	3,364 (28.9)	650 (22.6)	3,885 (23.3)	396 (21.6)	418 (21.3)
Ethnicity, n (%)							
Hispanic	11,932 (20.7)	3,942 (17.3)	3,378 (29.0)	646 (22.5)	3,298 (19.8)	332 (18.1)	336 (17.1)
Non-Hispanic	38,988 (67.5)	15,485 (68.0)	7,121 (61.2)	1,987 (69.1)	11,648 (69.9)	1,294 (70.4)	1,453 (74.1)
Missing	6,828 (11.8)	3,361 (14.7)	1,134 (9.7)	244 (8.5)	1,707 (10.3)	211 (11.5)	171 (8.7)
Smoking status, n (%)							
Current smoker	413 (0.7)	193 (0.8)	128 (1.1)	13 (0.5)	61 (0.4)	13 (0.7)	5 (0.3)
Previously smoked	740 (1.3)	417 (1.8)	111 (1.0)	27 (0.9)	145 (0.9)	25 (1.4)	15 (0.8)
Never smoked	2,831 (4.9)	1,468 (6.4)	750 (6.4)	121 (4.2)	404 (2.4)	50 (2.7)	38 (1.9)
Missing	53,764 (93.1)	20,710 (90.9)	10,644 (91.5)	2,716 (94.4)	16,043 (96.3)	1,749 (95.2)	1,902 (97.0)

Obese, n (%)*	10,952 (19.0)	4,905 (21.5)	1,366 (11.7)	580 (20.2)	3,246 (19.5)	406 (22.1)	449 (22.9)
Insurance, n (%)							
Commercial	29,145 (50.5)	13,134 (57.6)	5,672 (48.8)	1,482 (51.5)	7,243 (43.5)	758 (41.3)	856 (43.7)
Medicaid	8,652 (15.0)	2,341 (10.3)	2,223 (19.1)	542 (18.8)	2,891 (17.4)	312 (17.0)	343 (17.5)
Medicare	8,774 (15.2)	3,173 (13.9)	788 (6.8)	245 (8.5)	3,674 (22.1)	435 (23.7)	459 (23.4)
Other payor type	4,004 (6.9)	1,282 (5.6)	1,071 (9.2)	211 (7.3)	1,188 (7.1)	129 (7.0)	123 (6.3)
Uninsured	4,833 (8.4)	1,542 (6.8)	1,731 (14.9)	281 (9.8)	1,069 (6.4)	111 (6.0)	99 (5.1)
Missing	2,340 (4.1)	1,316 (5.8)	148 (1.3)	116 (4.0)	588 (3.5)	92 (5.0)	80 (4.1)
Month of COVID-19 diagnosis, n (%)							
Feb 2020	115 (0.2)	61 (0.3)	3 (<0.1)	4 (0.1)	34 (0.2)	5 (0.3)	8 (0.4)
Mar 2020	8,197 (14.2)	1,527 (6.7)	1,893 (16.3)	527 (18.3)	3,288 (19.7)	306 (16.7)	656 (33.5)
Apr 2020	18,591 (32.2)	6,018 (26.4)	3,480 (29.9)	926 (32.2)	6,684 (40.1)	676 (36.8)	807 (41.2)
May 2020	14,188 (24.6)	6,154 (27.0)	2,703 (23.2)	729 (25.3)	3,755 (22.5)	477 (26.0)	370 (18.9)
Jun 2020	14,832 (25.7)	7,846 (34.4)	3,073 (26.4)	597 (20.8)	2,826 (17.0)	371 (20.2)	119 (6.1)
Jul 2020	1,825 (3.2)	1,182 (5.2)	481 (4.1)	94 (3.3)	66 (0.4)	2 (0.1)	0 (0.0)
Mean weighted CCI (SD)	1.20 (2.06)	1.09 (2.03)	0.56 (1.32)	0.88 (1.76)	1.64 (2.31)	2.15 (2.46)	2.13 (2.39)

CCI, Charlson Comorbidity Index; COVID-19, coronavirus disease 2019; ER, emergency room; ICU, intensive care unit; SD, standard deviation

*No patient data were missing for obesity

Table 2 Long-term outcomes that were diagnosed >30–≤90 days or >90–≤180 days post COVID-19 by sub-cohort

Condition, n (%)	1. Outpatient (N=22,788)		2. ER on diagnosis (N=11,633)		3. ER (N=2,877)		4. Hospitalization without ICU (N=16,653)		5. ICU without ventilation (N=1,837)		6. ICU with ventilation (N=1,960)	
	>30–≤90 days	>90–≤180 days	>30–≤90 days	>90–≤180 days	>30–≤90 days	>90–≤180 days	>30–≤90 days	>90–≤180 days	>30–≤90 days	>90–≤180 days	>30–≤90 days	>90–≤180 days
Respiratory	1,846 (**8.1**)	1,846 (**8.1**)	520 (**4.5**)	484 (**4.2**)	322 (**11.2**)	311 (**10.8**)	2,124 (**12.8**)	1,599 (**9.6**)	296 (**16.1**)	222 (**12.1**)	547 (**27.9**)	381 (**19.4**)
CV	439 (**1.9**)	475 (**2.1**)	81 (**0.7**)	92 (**0.8**)	54 (**1.9**)	57 (**2.0**)	585 (**3.5**)	618 (**3.7**)	102 (**5.6**)	99 (**5.4**)	128 (**6.5**)	132 (**6.7**)
Mental health	940 (**4.1**)	1,081 (**4.7**)	231 (**2.0**)	293 (**2.5**)	144 (**5.0**)	154 (**5.4**)	819 (**4.9**)	866 (**5.2**)	119 (**6.5**)	118 (**6.4**)	167 (**8.5**)	144 (**7.3**)
Cancer	137 (**0.6**)	151 (**0.7**)	24 (**0.2**)	30 (**0.3**)	18 (**0.6**)	14 (**0.5**)	95 (**0.6**)	106 (**0.6**)	9 (**0.5**)	12 (**0.7**)	22 (**1.1**)	14 (**0.7**)
Respiratory and CV	131 (**0.6**)	124 (**0.5**)	16 (**0.1**)	30 (**0.3**)	17 (**0.6**)	19 (**0.7**)	201 (**1.2**)	181 (**1.1**)	30 (**1.6**)	26 (**1.4**)	63 (**3.2**)	47 (**2.4**)
Respiratory and mental health	205 (**0.9**)	216 (**0.9**)	65 (**0.6**)	63 (**0.5**)	48 (**1.7**)	33 (**1.1**)	262 (**1.6**)	245 (**1.5**)	45 (**2.4**)	34 (**1.9**)	83 (**4.2**)	50 (**2.6**)
Mental health and CV	53 (**0.2**)	49 (**0.2**)	8 (**0.1**)	9 (**0.1**)	4 (**0.1**)	2 (**0.1**)	57 (**0.3**)	64 (**0.4**)	10 (**0.5**)	11 (**0.6**)	7 (**0.4**)	11 (**0.6**)
Respiratory, CV, and mental health	57 (**0.3**)	52 (**0.2**)	9 (**0.1**)	10 (**0.1**)	6 (**0.2**)	9 (**0.3**)	98 (**0.6**)	84 (**0.5**)	17 (**0.9**)	15 (**0.8**)	30 (**1.5**)	31 (**1.6**)
No new conditions* (respiratory, CV, or mental health)	20,066 (**88.1**)	19,879 (**87.2**)	10,908 (**93.8**)	10,886 (**93.6**)	2,438 (**84.7**)	2,427 (**84.4**)	13,841 (**83.1**)	14,228 (**85.4**)	1,439 (**78.3**)	1,499 (**81.6**)	1,331 (**67.9**)	1,473 (**75.2**)

COVID-19, coronavirus disease 2019; CV, cardiovascular; ER, emergency room; ICU, intensive care unit

*Only conditions that appeared >30–≤180 days after COVID-19 diagnosis or hospital discharge are included; pre-existing conditions are excluded

Figure 1

Title: Overall study design.

Abbreviations: COVID-19, coronavirus disease 2019

Figure 2

Title: Relative risk of new respiratory conditions occurring from >30–≤180 days after COVID-19 diagnosis or hospital discharge.

Abbreviations: CCI, Charlson Comorbidity Index; COVID-19, coronavirus disease 2019; ER, emergency room; ICU, intensive care unit

Legend: Relative risk of new respiratory conditions occurring at (A) >30–≤90 days and (B) >90–≤180 days after COVID-19 diagnosis or hospital discharge. Graphs represent relative risk and 95% confidence intervals. Reference groups are: <18 years (age), female (sex), African American (race), Hispanic (ethnicity), non-obese (obesity), diagnosis in February and March 2020 (diagnosis month); commercial (insurance), cohort 1: outpatient (cohort).

Figure 3

Title: Relative risk of new cardiovascular conditions occurring from >30–≤180 days after COVID-19 diagnosis or hospital discharge.

Abbreviations: CCI, Charlson Comorbidity Index; COVID-19, coronavirus disease 2019; ER, emergency room; ICU, intensive care unit

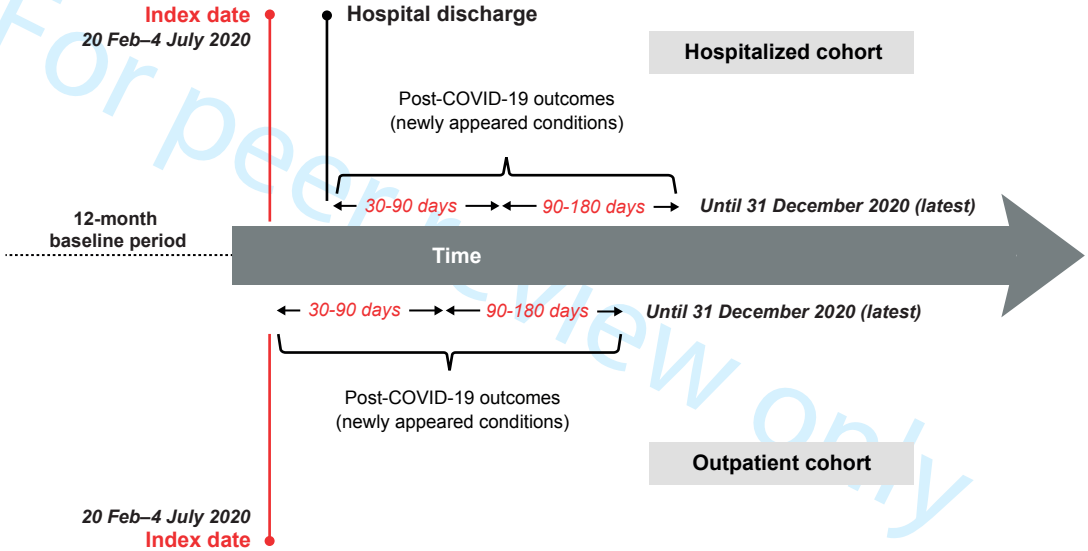
Legend: Relative risk of new cardiovascular conditions occurring at (A) >30–≤90 days and (B) >90–≤180 days after COVID-19 diagnosis or hospital discharge. Graphs represent relative risk and 95% confidence intervals. Reference groups are: <18 years (age), female (sex), African American (race), Hispanic (ethnicity), non-obese (obesity), diagnosis in February and March 2020 (diagnosis month); commercial (insurance), cohort 1: outpatient (cohort). Relative risk in the >30–≤90 days time window was not calculated as no new diagnoses were made in the reference group (<18 years) during this time.

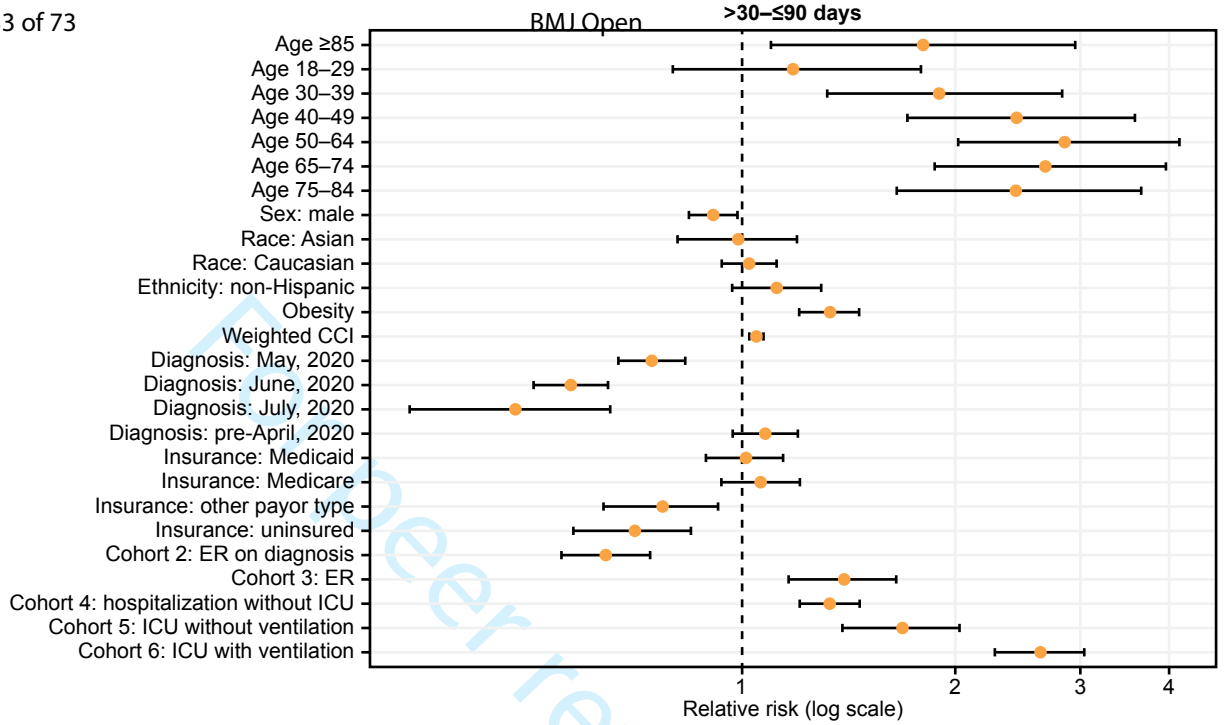
Figure 4

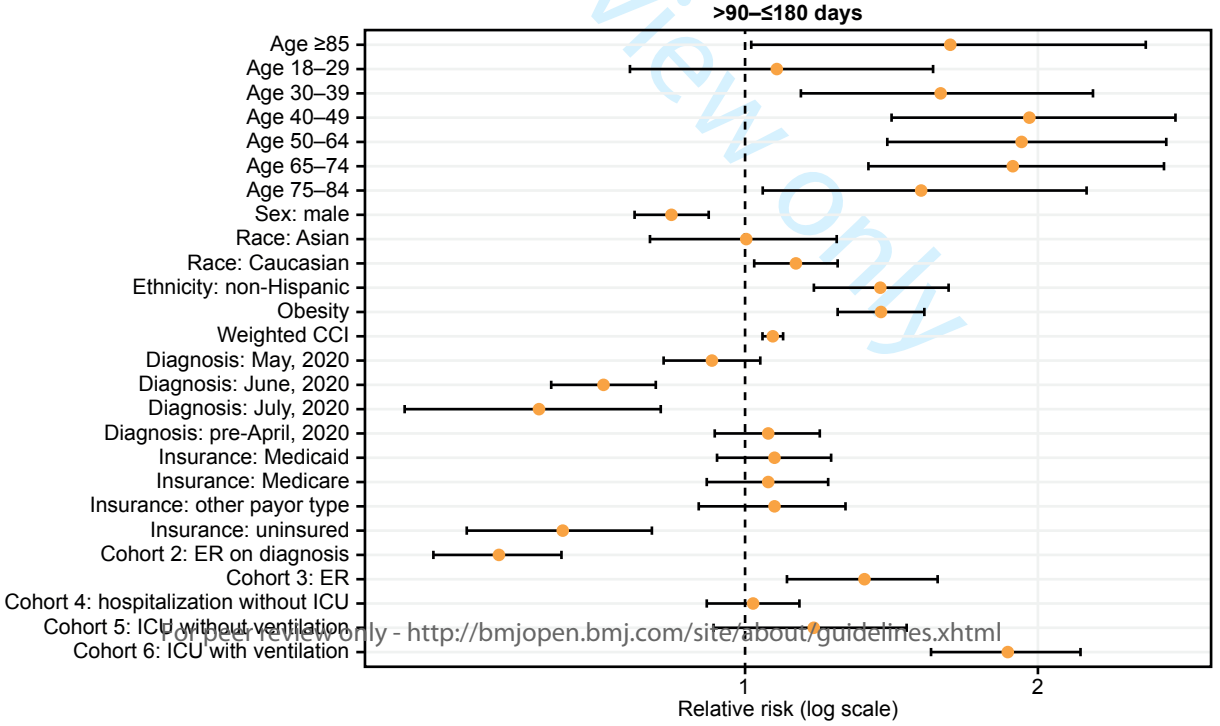
Title: Relative risk of new mental health conditions occurring from >30–≤180 days after COVID-19 diagnosis or hospital discharge.

Abbreviations: CCI, Charlson Comorbidity Index; COVID-19, coronavirus disease 2019; ER, emergency room; ICU, intensive care unit

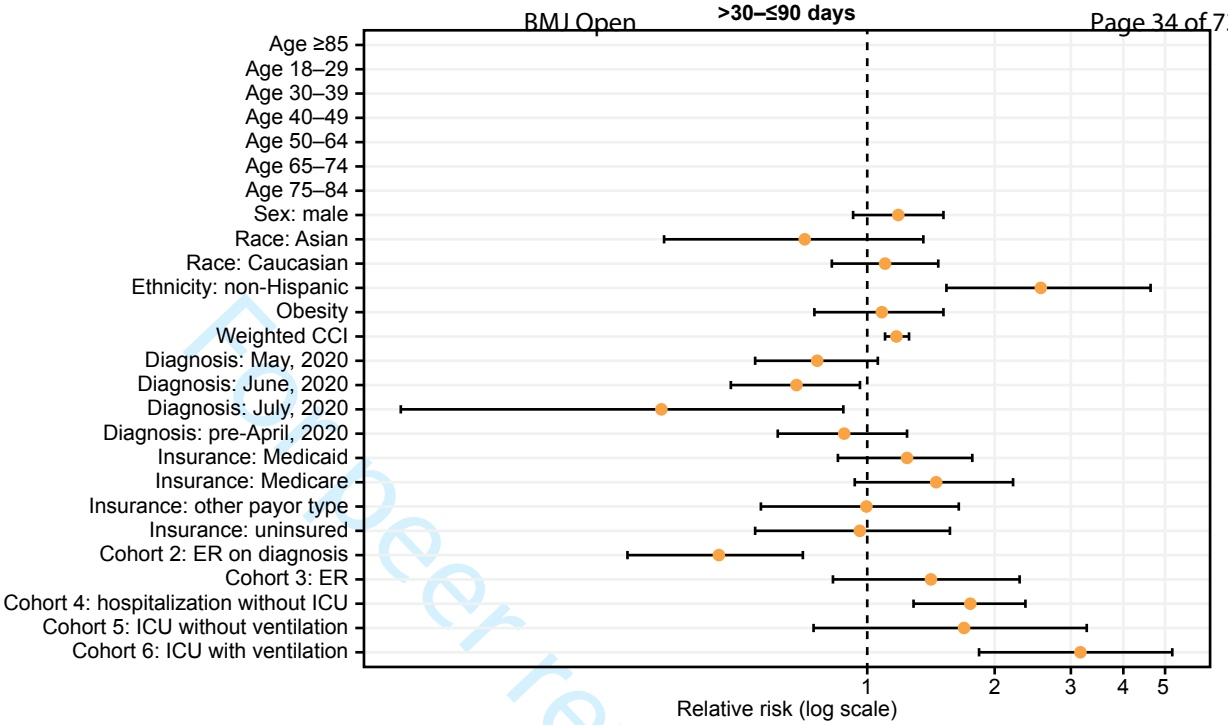
Legend: Relative risk of new mental health conditions occurring at (A) >30–≤90 days and (B) >90–≤180 days after COVID-19 diagnosis or hospital discharge. Graphs represent relative risk and 95% confidence intervals. Reference groups are: <18 years (age), female (sex), African American (race), Hispanic (ethnicity), non-obese (obesity), diagnosis in February and March 2020 (diagnosis month); commercial (insurance), cohort 1: outpatient (cohort).



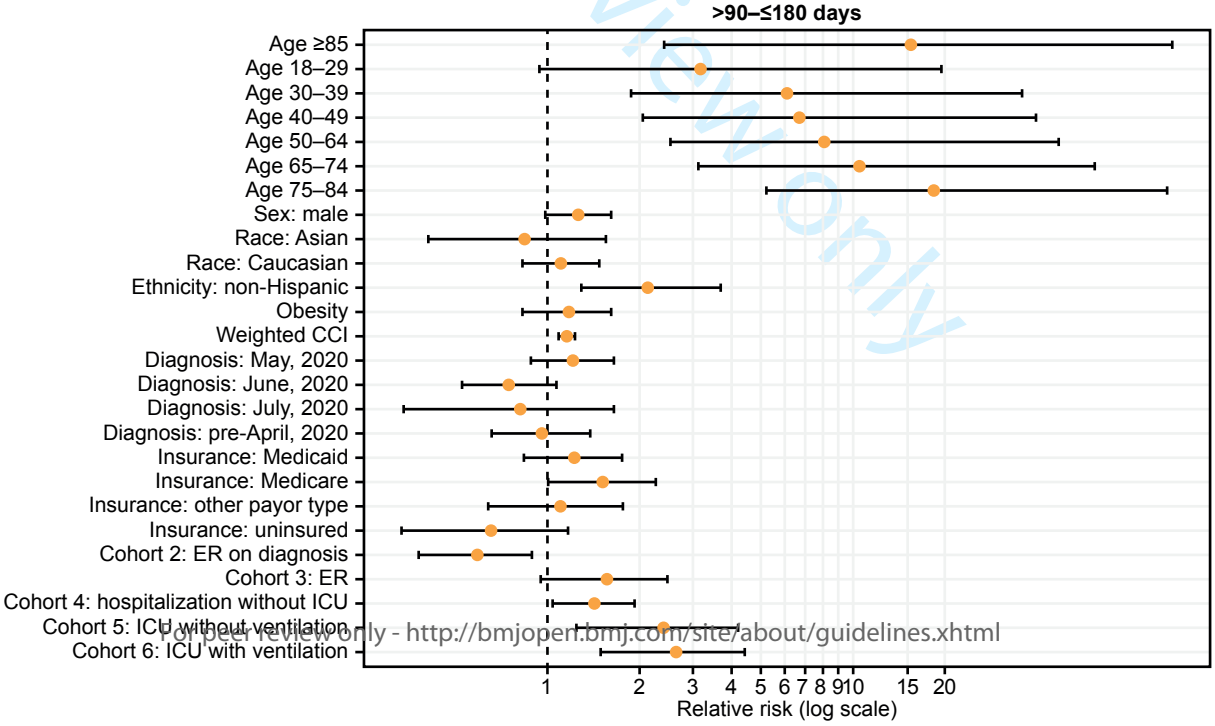


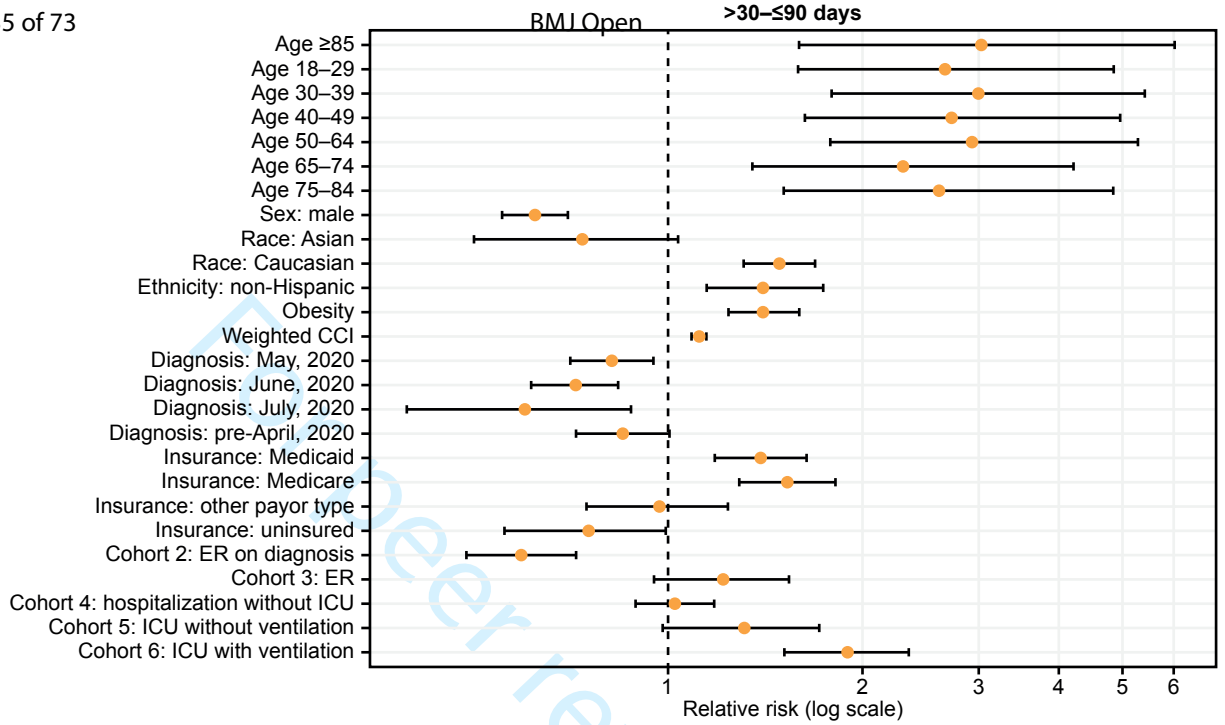


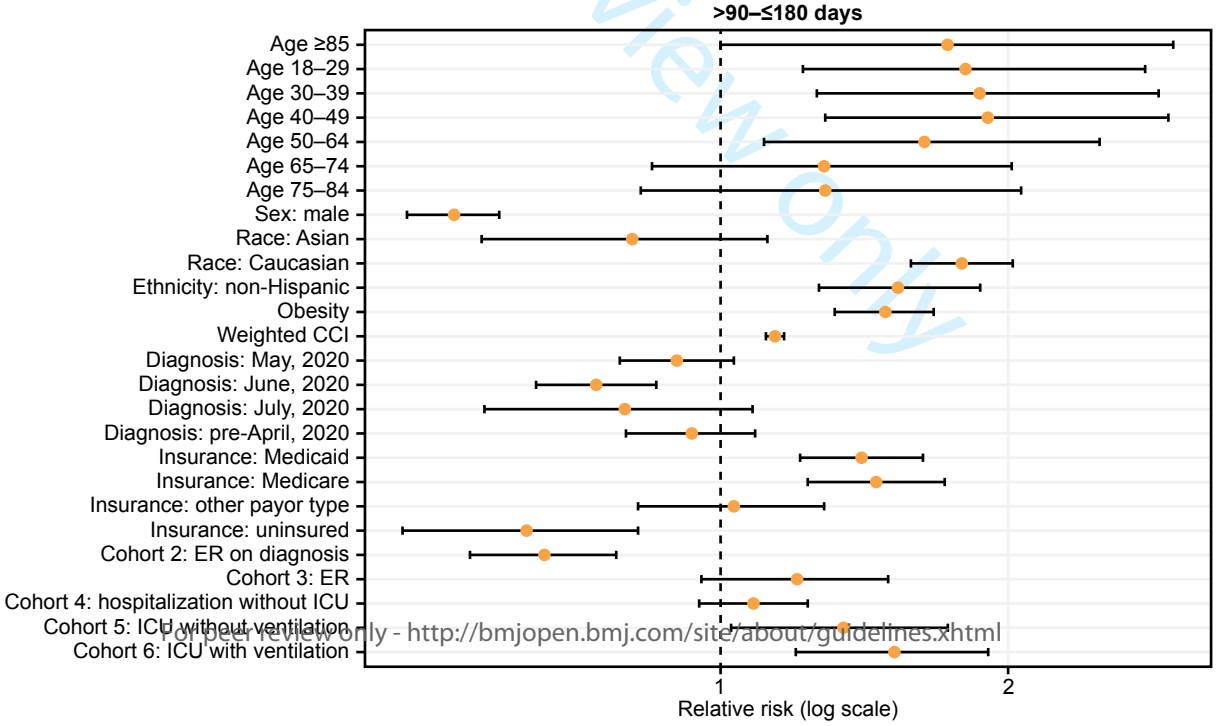
A



B



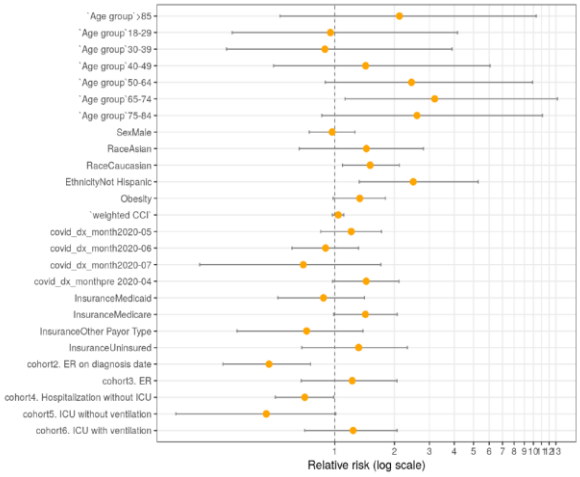




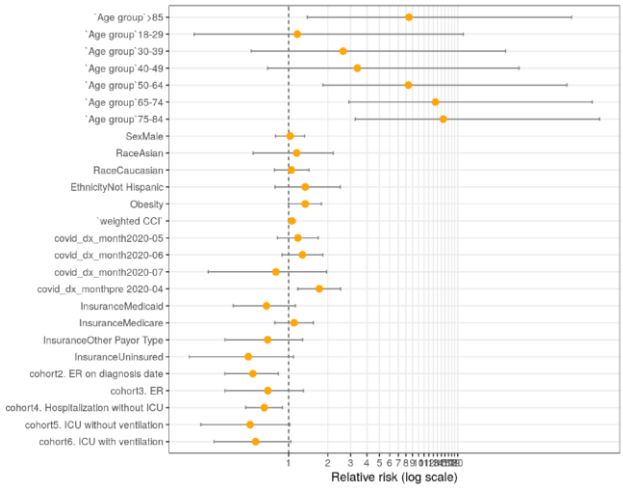
Supplemental materials

Supplemental Figure 1 Relative risk of a new cancer diagnosis from >30–≤180 days after COVID-19 diagnosis or hospital discharge.

A



B



CCI, Charlson Comorbidity Index; COVID-19, coronavirus disease 2019; ER, emergency room; ICU, intensive care unit

Relative risk of a new cancer diagnosis at (A) >30–≤90 days and (B) >90–≤180 days after COVID-19 diagnosis or hospital discharge. Graphs represent relative risk and 95% confidence intervals. Reference groups are: <18 years (age), female (sex), African American (race), Hispanic (ethnicity), non-obese (obesity), diagnosis in February and March 2020 (diagnosis month); commercial (insurance), cohort 1: outpatient (cohort).

Supplemental Table 1 List of the long-term outcomes studied and their classification

Long-term outcome
Respiratory
Asthma
Bronchiectasis
Bronchitis
COPD
Dyspnea
Emphysema
Influenza
Interstitial lung disease (fibrosis)
Pneumonia
Respiratory failure
Cardiovascular
Cardiac arrhythmia
Myocardial infarction
Pulmonary embolism
Pulmonary hypertension
Stroke
Mental health
Anxiety
Confusion or disorientation
Dementia
Depression
Encephalopathy
Memory loss

COPD, chronic obstructive pulmonary disease

Supplemental Table 2 List of comorbidities included in the Charlson Comorbidity Index

Comorbidity	
AIDS	Metastatic solid tumor
Cancer	Mild liver disease
Cerebrovascular disease	Moderate or severe liver disease
Chronic pulmonary disease	Moderate or severe renal disease
Congestive heart failure	Myocardial infarction
Dementia	Peptic ulcer disease
Diabetes with complication	Peripheral vascular disease
Diabetes without complication	Rheumatics
Hemiplegia	

AIDS, acquired immunodeficiency syndrome

Supplemental Table 3 Full baseline characteristics, COVID-19-related outcomes, symptoms, and tests

	All patients (N=57,748)	1. Outpatient (n=22,788)	2. ER on diagnosis (n=11,633)	3. ER (n=2,877)	4. Hospitalization without ICU (N=16,653)	5. ICU without ventilation (N=1,837)	6. ICU with ventilation (N=1,960)
Mean age (SD), years	47.93 (18.76)	46.78 (18.90)	40.95 (16.89)	44.00 (16.61)	53.17 (18.50)	55.92 (17.32)	57.70 (14.78)
Age group, n (%)							
<18 years	2,184 (3.8)	1,033 (4.5)	641 (5.5)	78 (2.7)	366 (2.2)	46 (2.5)	20 (1.0)
18–29 years	8,509 (14.7)	3,693 (16.2)	2,543 (21.9)	538 (18.7)	1574 (9.5)	89 (4.8)	72 (3.7)
30–39 years	8,972 (15.5)	3,541 (15.5)	2,496 (21.5)	579 (20.1)	2,046 (12.3)	175 (9.5)	135 (6.9)
40–49 years	9,362 (16.2)	3,664 (16.1)	2,205 (19.0)	555 (19.3)	2,442 (14.7)	253 (13.8)	243 (12.4)
50–64 years	16,103 (27.9)	6,231 (27.3)	2,706 (23.3)	780 (27.1)	4,937 (29.6)	626 (34.1)	823 (42.0)
65–74 years	7,065 (12.2)	2,658 (11.7)	689 (5.9)	230 (8.0)	2,683 (16.1)	363 (19.8)	442 (22.6)
75–84 years	3,620 (6.3)	1,279 (5.6)	239 (2.1)	76 (2.6)	1,647 (9.9)	195 (10.6)	184 (9.4)
≥85 years	891 (1.5)	303 (1.3)	54 (0.5)	22 (0.8)	440 (2.6)	44 (2.4)	28 (1.4)

Missing	1,042 (1.8)	386 (1.7)	60 (0.5)	19 (0.7)	518 (3.1)	46 (2.5)	13 (0.7)
Sex, n (%)							
Female	30,782 (53.3)	12,856 (56.4)	6,115 (52.6)	1,721 (59.8)	8,487 (51.0)	829 (45.1)	774 (39.5)
Male	26,939 (46.6)	9,920 (43.5)	5,515 (47.4)	1,152 (40.0)	8,160 (49.0)	1,008 (54.9)	1,184 (60.4)
Missing	27 (<0.1)	12 (<0.1)	3 (<0.1)	4 (<0.1)	6 (<0.1)	0 (0.0)	2 (<0.1)
Race, n (%)							
African American	13,183 (22.8)	3,473 (15.2)	3,178 (27.3)	790 (27.5)	4,675 (28.1)	551 (30.0)	516 (26.3)
Asian	1,848 (3.2)	639 (2.8)	438 (3.8)	100 (3.5)	555 (3.3)	41 (2.2)	75 (3.8)
Caucasian	29,074 (50.3)	13,746 (60.3)	4,653 (40.0)	1,337 (46.5)	7,538 (45.3)	849 (46.2)	951 (48.5)
Missing	13,643 (23.6)	4,930 (21.6)	3,364 (28.9)	650 (22.6)	3,885 (23.3)	396 (21.6)	418 (21.3)
Ethnicity, n (%)							
Hispanic	11,932 (20.7)	3,942 (17.3)	3,378 (29.0)	646 (22.5)	3,298 (19.8)	332 (18.1)	336 (17.1)
Non-Hispanic	38,988 (67.5)	15,485 (68.0)	7,121 (61.2)	1,987 (69.1)	11,648 (69.9)	1,294 (70.4)	1,453 (74.1)
Missing	6,828 (11.8)	3,361 (14.7)	1,134 (9.7)	244 (8.5)	1,707 (10.3)	211 (11.5)	171 (8.7)

Region, n (%)							
Midwest	22,133 (38.3)	8,137 (35.7)	5,250 (45.1)	1,364 (47.4)	5,686 (34.1)	830 (45.2)	866 (44.2)
Northwest	20,671 (35.8)	8,018 (35.2)	3,261 (28.0)	875 (30.4)	7,375 (44.3)	496 (27.0)	646 (33.0)
South	8,548 (14.8)	3,463 (15.2)	2,004 (17.2)	367 (12.8)	2,212 (13.3)	271 (14.8)	231 (11.8)
West	4,430 (7.7)	2,379 (10.4)	673 (5.8)	169 (5.9)	873 (5.2)	178 (9.7)	158 (8.1)
Missing	1,966 (3.4)	791 (3.5)	445 (3.8)	102 (3.5)	507 (3.0)	62 (3.4)	59 (3.0)
Division, n (%)							
East North Central	15,381 (26.6)	4,833 (21.2)	3,822 (32.9)	982 (34.1)	4,536 (27.2)	551 (30.0)	657 (33.5)
East South Central	1,769 (3.1)	664 (2.9)	404 (3.5)	62 (2.2)	472 (2.8)	124 (6.8)	43 (2.2)
Middle Atlantic	15,163 (26.3)	6,516 (28.6)	1,718 (14.8)	527 (18.3)	5,622 (33.8)	338 (18.4)	442 (22.6)
Mountain	2,221 (3.8)	1,281 (5.6)	257 (2.2)	48 (1.7)	457 (2.7)	78 (4.2)	100 (5.1)
New England	5,478 (9.5)	1,491 (6.5)	1,538 (13.2)	345 (12.0)	1,744 (10.5)	158 (8.6)	202 (10.3)
Other/unknown	2,085 (3.6)	847 (3.7)	467 (4.0)	112 (3.9)	532 (3.2)	64 (3.5)	63 (3.2)
Pacific	2,205 (3.8)	1,096 (4.8)	416 (3.6)	121 (4.2)	414 (2.5)	100 (5.4)	58 (3.0)

South Atlantic/ West South Central	6,767 (11.7)	2,792 (12.3)	1,599 (13.7)	303 (10.5)	1,738 (10.4)	147 (8.0)	188 (9.6)
West North Central	6,679 (11.6)	3,268 (14.3)	1,412 (12.1)	377 (13.1)	1,138 (6.8)	277 (15.1)	207 (10.6)
Smoking status, n (%)							
Current smoker	413 (0.7)	193 (0.8)	128 (1.1)	13 (0.5)	61 (0.4)	13 (0.7)	5 (0.3)
Previously smoked	740 (1.3)	417 (1.8)	111 (1.0)	27 (0.9)	145 (0.9)	25 (1.4)	15 (0.8)
Never smoked	2,831 (4.9)	1,468 (6.4)	750 (6.4)	121 (4.2)	404 (2.4)	50 (2.7)	38 (1.9)
Missing	53,764 (93.1)	20,710 (90.9)	10,644 (91.5)	2,716 (94.4)	16,043 (96.3)	1,749 (95.2)	1,902 (97.0)
Obese, n (%)							
No	47,796 (81.0)	17,883 (78.5)	10,267 (88.3)	2,297 (79.8)	13,407 (80.5)	1,431 (77.9)	1,511 (77.1)
Yes	10,952 (19.0)	4,905 (21.5)	1,366 (11.7)	580 (20.2)	3,246 (19.5)	406 (22.1)	449 (22.9)
Pregnant n (%)							
No	56,137 (97.2)	22,246 (97.6)	11,409 (98.1)	2,802 (97.4)	15,931 (95.7)	1,813 (98.7)	1,936 (98.8)
Yes	1,611 (2.8)	542 (2.4)	224 (1.9)	75 (2.6)	722 (4.3)	24 (1.3)	24.2 (1.2)
Insurance, n (%)							

Commercial	29,145 (50.5)	13,134 (57.6)	5,672 (48.8)	1,482 (51.5)	7,243 (43.5)	758 (41.3)	856 (43.7)
Medicaid	8,652 (15.0)	2,341 (10.3)	2,223 (19.1)	542 (18.8)	2,891 (17.4)	312 (17.0)	343 (17.5)
Medicare	8,774 (15.2)	3,173 (13.9)	788 (6.8)	245 (8.5)	3674 (22.1)	435 (23.7)	459 (23.4)
Other payor type	4,004 (6.9)	1,282 (5.6)	1,071 (9.2)	211 (7.3)	1,188 (7.1)	129 (7.0)	123 (6.3)
Uninsured	4,833 (8.4)	1,542 (6.8)	1,731 (14.9)	281 (9.8)	1,069 (6.4)	111 (6.0)	99 (5.1)
Missing	2,340 (4.1)	1,316 (5.8)	148 (1.3)	116 (4.0)	588 (3.5)	92 (5.0)	80 (4.1)
Month of COVID-19 diagnosis, n (%)							
Feb 2020	115 (0.2)	61 (0.3)	3 (<0.1)	4 (0.1)	34 (0.2)	5 (0.3)	8 (0.4)
Mar 2020	8,197 (14.2)	1,527 (6.7)	1,893 (16.3)	527 (18.3)	3,288 (19.7)	306 (16.7)	656 (33.5)
Apr 2020	18,591 (32.2)	6,018 (26.4)	3,480 (29.9)	926 (32.2)	6,684 (40.1)	676 (36.8)	807 (41.2)
May 2020	14,188 (24.6)	6,154 (27.0)	2,703 (23.2)	729 (25.3)	3755 (22.5)	477 (26.0)	370 (18.9)
Jun 2020	14,832 (25.7)	7,846 (34.4)	3,073 (26.4)	597 (20.8)	2,826 (17.0)	371 (20.2)	119 (6.1)
Jul 2020	1,825 (3.2)	1,182 (5.2)	481 (4.1)	94 (3.3)	66 (0.4)	2 (0.1)	0 (0.0)
Diabetes with complication, n (%)							

No	53,804 (93.2)	21,658 (95.0)	11,356 (97.6)	2,774 (96.4)	14,812 (88.9)	1,553 (84.5)	1,651 (84.2)
Yes	3,944 (6.8)	1,130 (5.0)	277 (2.4)	103 (3.6)	1,841 (11.1)	284 (15.5)	309 (15.8)
Congestive heart failure, n (%)							
No	54,048 (93.6)	21,702 (95.2)	11,430 (98.3)	2,800 (97.3)	14,938 (89.7)	1,553 (84.5)	1,625 (82.9)
Yes	3,700 (6.4)	1,086 (4.8)	203 (1.7)	77 (2.7)	1,715 (10.3)	284 (15.5)	335 (17.1)
Cerebrovascular disease, n (%)							
No	55,258 (95.7)	21,957 (96.4)	11,485 (98.7)	2,811 (97.7)	15,547 (93.4)	1,662 (90.5)	1,796 (91.6)
Yes	2,490 (4.3)	831 (3.6)	148 (1.3)	66 (2.3)	1,106 (6.6)	175 (9.5)	164 (8.4)
Moderate or severe renal disease, n (%)							
No	53,066 (91.9)	21,454 (94.1)	11,387 (97.9)	2,759 (95.9)	14,387 (86.4)	1,497 (81.5)	1,582 (80.7)
Yes	4,682 (8.1)	1,334 (5.9)	246 (2.1)	118 (4.1)	2,266 (13.6)	340 (18.5)	378 (19.3)
Diabetes without complication, n (%)							
No	47,489 (82.2)	19,807 (86.9)	10,435 (89.7)	2,522 (87.7)	12,313 (73.9)	1,184 (64.5)	1,228 (62.7)

Yes	10,259 (17.8)	2,981 (13.1)	1,198 (10.3)	355 (12.3)	4,340 (26.1)	653 (35.5)	732 (37.3)
Chronic pulmonary disease, n (%)							
No	47,794 (82.8)	19,225 (84.4)	10,203 (87.7)	2,359 (82.0)	13,145 (78.9)	1,439 (78.3)	1,423 (72.6)
Yes	9,954 (17.2)	3,563 (15.6)	1,430 (12.3)	518 (18.0)	3,508 (21.1)	398 (21.7)	537 (27.4)
Mild liver disease, n (%)							
No	55,817 (96.7)	22,095 (97.0)	11,441 (98.3)	2,788 (96.9)	15,890 (95.4)	1,746 (95.0)	1,857 (94.7)
Yes	1,931 (3.3)	693 (3.0)	192 (1.7)	89 (3.1)	763 (4.6)	91 (5.0)	103 (5.3)
Peripheral vascular disease, n (%)							
No	55,049 (95.3)	21,773 (95.5)	11,461 (98.5)	2,808 (97.6)	15,516 (93.2)	1,674 (91.1)	1,817 (92.7)
Yes	2,699 (4.7)	1,015 (4.5)	172 (1.5)	69 (2.4)	1,137 (6.8)	163 (8.9)	143 (7.3)
Cancer, n (%)							
No	53,687 (93.0)	20,822 (91.4)	11,198 (96.3)	2,686 (93.4)	15,465 (92.9)	1,689 (91.9)	1,827 (93.2)
Yes	4,061 (7.0)	1,966 (8.6)	435 (3.7)	191 (6.6)	1,188 (7.1)	148 (8.1)	133 (6.8)

Myocardial infarction, n (%)							
No	55,528 (96.2)	22,273 (97.7)	11,459 (98.5)	2,828 (98.3)	15,528 (93.2)	1,659 (90.3)	1,781 (90.9)
Yes	2,220 (3.8)	515 (2.3)	174 (1.5)	49 (1.7)	1,125 (6.8)	178 (9.7)	179 (9.1)
Dementia, n (%)							
No	55,833 (96.7)	22,213 (97.5)	11,541 (99.2)	2,843 (98.8)	15,663 (94.1)	1,697 (92.4)	1,876 (95.7)
Yes	1,915 (3.3)	575 (2.5)	92 (0.8)	34 (1.2)	990 (5.9)	140 (7.6)	84 (4.3)
Peptic ulcer disease, n (%)							
No	57,305 (99.2)	22,620 (99.3)	11,599 (99.7)	2,864 (99.5)	16,494 (99.0)	1,812 (98.6)	1,916 (97.8)
Yes	443 (0.8)	168 (0.7)	34 (0.3)	13 (0.5)	159 (1.0)	25 (1.4)	44 (2.2)
Hemiplegia, n (%)							
No	57,192 (99.0)	22,647 (99.4)	11,596 (99.7)	2,870 (99.8)	16,402 (98.5)	1,783 (97.1)	1,894 (96.6)
Yes	556 (1.0)	141 (0.6)	37 (0.3)	7 (0.2)	251 (1.5)	54 (2.9)	66 (3.4)

Rheumatics, n (%)							
No	56,674 (98.1)	22,348 (98.1)	11,515 (99.0)	2,831 (98.4)	16,285 (97.8)	1,790 (97.4)	1,905 (97.2)
Yes	1,074 (1.9)	440 (1.9)	118 (1.0)	46 (1.6)	368 (2.2)	47 (2.6)	55 (2.8)
Metastatic solid tumor, n (%)							
No	57,146 (99.0)	22,460 (98.6)	11,601 (99.7)	2,858 (99.3)	16,472 (98.9)	1,810 (98.5)	1,945 (99.2)
Yes	602 (1.0)	328 (1.4)	32 (0.3)	19 (0.7)	181 (1.1)	27 (1.5)	15 (0.8)
Moderate or severe liver disease, n (%)							
No	57,495 (99.6)	22,703 (99.6)	11,625 (99.9)	2,872 (99.8)	16,547 (99.4)	1,818 (99.0)	1,930 (98.5)
Yes	253 (0.4)	85 (0.4)	8 (0.1)	5 (0.2)	106 (0.6)	19 (1.0)	30 (1.5)
AIDS, n (%)							
No	56,640 (98.1)	22,229 (97.5)	11,500 (98.9)	2,813 (97.8)	1,6376 (98.3)	1,794 (97.7)	1,928 (98.4)
Yes	1,108 (1.9)	559 (2.5)	133 (1.1)	64 (2.2)	277 (1.7)	43 (2.3)	32 (1.6)
Mean weighted CCI (SD)	1.20 (2.06)	1.09 (2.03)	0.56 (1.32)	0.88 (1.76)	1.64 (2.31)	2.15 (2.46)	2.13 (2.39)

COPD, n (%)							
No	54,029 (93.6)	21,688 (95.2)	11,382 (97.8)	2,770 (96.3)	14,961 (89.8)	1,594 (86.8)	1,634 (83.4)
Yes	3,719 (6.4)	1,100 (4.8)	251 (2.2)	107 (3.7)	1,692 (10.2)	243 (13.2)	326 (16.6)
Diabetes, n (%)							
No	46,563 (80.6)	19,539 (85.7)	10,380 (89.2)	2,489 (86.5)	11,878 (71.3)	1,119 (60.9)	1,158 (59.1)
Yes	11,185 (19.4)	3,249 (14.3)	1,253 (10.8)	388 (13.5)	4,775 (28.7)	718 (39.1)	802 (40.9)
Hypertension, n (%)							
No	37,710 (65.3)	16,023 (70.3)	9,335 (80.2)	2,130 (74.0)	8,643 (51.9)	793 (43.2)	786 (40.1)
Yes	20,038 (34.7)	6,765 (29.7)	2,298 (19.8)	747 (26.0)	8,010 (48.1)	1,044 (56.8)	1,174 (59.9)
Asthma, n (%)							
No	51,726 (89.6)	20,663 (90.7)	10,587 (91.0)	2,517 (87.5)	14,604 (87.7)	1,652 (89.9)	1,703 (86.9)
Yes	6,022 (10.4)	2,125 (9.3)	1,046 (9.0)	360 (12.5)	2,049 (12.3)	185 (10.1)	257 (13.1)
Chronic renal disease, n (%)							

No	53,421 (92.5)	21,585 (94.7)	11,411 (98.1)	2,768 (96.2)	14,520 (87.2)	1,524 (83.0)	1,613 (82.3)
Yes	4,327 (7.5)	1,203 (5.3)	222 (1.9)	109 (3.8)	2,133 (12.8)	313 (17.0)	347 (17.7)
Other chronic respiratory disease, n (%)							
No	55,858 (96.7)	21,900 (96.1)	11,365 (97.7)	2,738 (95.2)	16,186 (97.2)	1,779 (96.8)	1,890 (96.4)
Yes	1,890 (3.3)	888 (3.9)	268 (2.3)	139 (4.8)	467 (2.8)	58 (3.2)	70 (3.6)
Chronic ischemic heart disease, n (%)							
No	52,308 (90.6)	20,978 (92.1)	11,270 (96.9)	2,733 (95.0)	14,236 (85.5)	1,491 (81.2)	1,600 (81.6)
Yes	5,440 (9.4)	1,810 (7.9)	363 (3.1)	144 (5.0)	2,417 (14.5)	346 (18.8)	360 (18.4)
End stage renal disease, n (%)							
No	56,530 (97.9)	22,463 (98.6)	11,560 (99.4)	2,849 (99.0)	16,079 (96.6)	1,740 (94.7)	1,839 (93.8)
Yes	1,218 (2.1)	325 (1.4)	73 (0.6)	28 (1.0)	574 (3.4)	97 (5.3)	121 (6.2)
Liver disease							
No	55,348 (95.8)	21,961 (96.4)	11,407 (98.1)	2,775 (96.5)	15,683 (94.2)	1,718 (93.5)	1,804 (92.0)

Yes	2,400 (4.2)	827 (3.6)	226 (1.9)	102 (3.5)	970 (5.8)	119 (6.5)	156 (8.0)
HIV, n (%)							
No	57,000 (98.7)	22,425 (98.4)	11,533 (99.1)	2,836 (98.6)	16,456 (98.8)	1,811 (98.6)	1,939 (98.9)
Yes	748 (1.3)	363 (1.6)	100 (0.9)	41 (1.4)	197 (1.2)	26 (1.4)	21 (1.1)
Immunocompromised, n (%)							
No	53,530 (92.7)	21,906 (96.1)	11,403 (98.0)	2,802 (97.4)	14,429 (86.6)	1,490 (81.1)	1,500 (76.5)
Yes	4,218 (7.3)	882 (3.9)	230 (2.0)	75 (2.6)	2,224 (13.4)	347 (18.9)	460 (23.5)

AIDS, acquired immunodeficiency syndrome; CCI, Charlson Comorbidity Index; COPD, chronic obstructive pulmonary disease; COVID-19, coronavirus disease 2019; ER, emergency room; HIV, human immunodeficiency virus; ICU, intensive care unit; SD, standard deviation

Supplemental Table 4 Full list of long-term outcomes that occurred >30–≤90 days or >90–≤180 days post COVID-19 diagnosis or hospitalization

	1. Outpatient (N=22,788)		2. ER on diagnosis date (N=11,633)		3. ER (N=2,877)		4. Hospitalization without ICU (N=16,653)		5. ICU without ventilation (N=1,837)		6. ICU with ventilation (N=1,960)	
	>30–≤90 days	>90–≤180 days	>30–≤90 days	>90–≤180 days	>30–≤90 days	>90–≤180 days	>30–≤90 days	>90–≤180 days	>30–≤90 days	>90–≤180 days	>30–≤90 days	>90–≤180 days
Pneumonia, n (%)												
No	22,368 (98.2)	22,582 (99.1)	11,520 (99.0)	11,579 (99.5)	2,804 (97.5)	2,848 (99.0)	15,612 (93.7)	16,187 (97.2)	1,689 (91.9)	1,768 (96.2)	1,687 (86.1)	1,801 (91.9)
Yes	420 (1.8)	206 (0.9)	113 (1.0)	54 (0.5)	73 (2.5)	29 (1.0)	1,041 (6.3)	466 (2.8)	148 (8.1)	69 (3.8)	273 (13.9)	159 (8.1)
Asthma, n (%)												
No	22,487 (98.7)	22,459 (98.6)	11,532 (99.1)	11,503 (98.9)	2,825 (98.2)	2,822 (98.1)	16,424 (98.6)	16,410 (98.5)	1,810 (98.5)	1,815 (98.8)	1,919 (97.9)	1,922 (98.1)
Yes	301 (1.3)	329 (1.4)	101 (0.9)	130 (1.1)	52 (1.8)	55 (1.9)	229 (1.4)	243 (1.5)	27 (1.5)	22 (1.2)	41 (2.1)	38 (1.9)
COPD, n (%)												

No	22,626 (99.3)	22,776 (99.9)	11,615 (99.8)	11,606 (99.8)	2,865 (99.6)	2,858 (99.3)	16,460 (98.8)	16,442 (98.7)	1,802 (98.1)	1,804 (98.2)	1,894 (96.6)	1,919 (97.9)
Yes	162 (0.7)	179 (0.8)	18 (0.2)	28 (0.2)	12 (0.4)	19 (0.7)	193 (1.2)	211 (1.3)	35 (1.9)	33 (1.8)	66 (3.4)	41 (2.1)
Influenza, n (%)												
No	22,783 (100.0)	22,776 (99.9)	11,630 (100.0)	11,631 (100.0)	2,875 (99.9)	2,877 (100.0)	16,646 (100.0)	16,648 (100.0)	1,837 (100.0)	1,837 (100.0)	1,960 (100.0)	1,960 (100.0)
Yes	5 (0.0)	12 (0.1)	3 (0.0)	2 (0.0)	2 (0.1)	0 (0.0)	7 (0.0)	5 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Stroke, n (%)												
No	22,695 (99.6)	22,696 (99.6)	11,619 (99.9)	11,611 (99.8)	2,865 (99.6)	2,865 (99.6)	16,535 (99.3)	16,506 (99.1)	1,813 (98.7)	1,811 (98.6)	1,935 (98.7)	1,926 (98.3)
Yes	93 (0.4)	92 (0.4)	14 (0.1)	22 (0.2)	12 (0.4)	12 (0.4)	118 (0.7)	147 (0.9)	24 (1.3)	26 (1.4)	25 (1.3)	34 (1.7)
Anxiety, n (%)												
No	22,250 (97.6)	22,169 (97.3)	11,491 (98.8)	11,451 (98.4)	2,779 (96.6)	2,774 (96.4)	16,274 (97.7)	16,268 (97.7)	1,793 (97.6)	1,794 (97.7)	1,889 (96.4)	1,896 (96.7)
Yes	538 (2.4)	619 (2.7)	142 (1.2)	182 (1.6)	98 (3.4)	103 (3.6)	379 (2.3)	385 (2.3)	44 (2.4)	43 (2.3)	71 (3.6)	64 (3.3)

Depression, n (%)												
No	22,456 (98.5)	22,361 (98.1)	11,535 (99.2)	11,502 (98.9)	2,833 (98.5)	2,822 (98.1)	16,375 (98.3)	16,314 (98.0)	1,789 (97.4)	1,786 (97.2)	1,907 (97.3)	1,908 (97.3)
Yes	332 (1.5)	427 (1.9)	98 (0.8)	131 (1.1)	44 (1.5)	55 (1.9)	278 (1.7)	339 (2.0)	48 (2.6)	51 (2.8)	53 (2.7)	52 (2.7)
Myocardial infarction, n (%)												
No	22,691 (99.6)	22,671 (99.5)	11,617 (99.9)	11,617 (99.9)	2,866 (99.6)	2,868 (99.7)	16,497 (99.1)	16,492 (99.0)	1,810 (98.5)	1,806 (98.3)	1,927 (98.3)	1,926 (98.3)
Yes	97 (0.4)	117 (0.5)	16 (0.1)	16 (0.1)	11 (0.4)	9 (0.3)	156 (0.9)	161 (1.0)	27 (1.5)	31 (1.7)	33 (1.7)	34 (1.7)
Interstitial lung disease (fibrosis), n (%)												
No	22,741 (99.8)	22,728 (99.7)	11,623 (99.9)	11,621 (99.9)	2,874 (99.9)	2,873 (99.9)	16,592 (99.6)	16,578 (99.5)	1,830 (99.6)	1,828 (99.5)	1,929 (98.4)	1,922 (98.1)
Yes	47 (0.2)	60 (0.3)	10 (0.1)	12 (0.1)	3 (0.1)	4 (0.1)	61 (0.4)	75 (0.5)	7 (0.4)	9 (0.5)	31 (1.6)	38 (1.9)
Dsypnea, n (%)												
No	21,660 (95.1)	21,567 (94.6)	11,311 (97.2)	11,329 (97.4)	2,675 (93.0)	2,649 (92.1)	15,781 (94.8)	15,838 (95.1)	1,720 (93.6)	1,723 (93.8)	1,759 (89.7)	1,783 (91.0)

Yes	1,128 (4.9)	1,221 (5.4)	322 (2.8)	304 (2.6)	202 (7.0)	228 (7.9)	872 (5.2)	815 (4.9)	117 (6.4)	114 (6.2)	201 (10.3)	177 (9.0)
Respiratory failure, n (%)												
No	22,614 (99.2)	22,654 (99.4)	11,606 (99.8)	11,609 (99.8)	2,863 (99.5)	2,868 (99.7)	16,199 (97.3)	16,380 (98.4)	1,757 (95.6)	1,780 (96.9)	1,685 (86.0)	1,801 (91.9)
Yes	174 (0.8)	134 (0.6)	27 (0.2)	24 (0.2)	14 (0.5)	9 (0.3)	454 (2.7)	273 (1.6)	80 (4.4)	57 (3.1)	275 (14.0)	159 (8.1)
Pulmonary hypertension, n (%)												
No	22,719 (99.7)	22,716 (99.7)	11,626 (99.9)	11,622 (99.9)	2,872 (99.8)	2,874 (99.9)	16,566 (99.5)	16,551 (99.4)	1,824 (99.3)	1,821 (99.1)	1,944 (99.2)	1,927 (98.3)
Yes	69 (0.3)	72 (0.3)	7 (0.1)	11 (0.1)	5 (0.2)	3 (0.1)	87 (0.5)	102 (0.6)	13 (0.7)	16 (0.9)	16 (0.8)	33 (1.7)
Pulmonary embolism, n (%)												
No	22,714 (99.7)	22,719 (99.7)	11,622 (99.9)	11,623 (99.9)	2,865 (99.6)	2,863 (99.5)	16,478 (98.9)	16,513 (99.2)	1,809 (98.5)	1,815 (98.8)	1,918 (97.9)	1,938 (98.9)
Yes	74 (0.3)	69 (0.3)	11 (0.1)	10 (0.1)	12 (0.4)	14 (0.5)	175 (1.1)	140 (0.8)	28 (1.5)	22 (1.2)	42 (2.1)	22 (1.1)
Bronchitis, n (%)												

No	22,707 (99.6)	22,705 (99.6)	11,611 (99.8)	11,619 (99.9)	2,862 (99.5)	2,868 (99.7)	16,583 (99.6)	16,598 (99.7)	1,830 (99.6)	1,832 (99.7)	1,939 (98.9)	1,940 (99.0)
Yes	81 (0.4)	83 (0.4)	22 (0.2)	14 (0.1)	15 (0.5)	9 (0.3)	70 (0.4)	55 (0.3)	7 (0.4)	5 (0.3)	21 (1.1)	20 (1.0)
Emphysema, n (%)												
No	22,727 (99.7)	22,722 (99.7)	11,626 (99.9)	11,620 (99.9)	2,872 (99.8)	2,870 (99.8)	16,591 (99.6)	16,577 (99.5)	1,815 (98.8)	1,822 (99.2)	1,941 (99.0)	1,944 (99.2)
Yes	61 (0.3)	66 (0.3)	7 (0.1)	13 (0.1)	5 (0.2)	7 (0.2)	62 (0.4)	76 (0.5)	22 (1.2)	15 (0.8)	19 (1.0)	16 (0.8)
Bronchiectasis, n (%)												
No	22,765 (99.9)	22,763 (99.9)	11,632 (100.0)	11,629 (100.0)	2,876 (100.0)	2,874 (99.9)	16,630 (99.9)	16,625 (99.8)	1,836 (99.9)	1,836 (99.9)	1,951 (99.5)	1,952 (99.6)
Yes	23 (0.1)	25 (0.1)	1 (0.0)	4 (0.0)	1 (0.0)	3 (0.1)	23 (0.1)	28 (0.2)	1 (0.1)	1 (0.1)	9 (0.5)	8 (0.4)
Encephalopathy, n (%)												
No	22,709 (99.7)	22,732 (99.8)	11,624 (99.9)	11,627 (99.9)	2,872 (99.8)	2,874 (99.9)	16,545 (99.4)	16,554 (99.4)	1,809 (98.5)	1,816 (98.9)	1,911 (97.5)	1,923 (98.1)
Yes	79 (0.3)	56 (0.2)	9 (0.1)	6 (0.1)	5 (0.2)	3 (0.1)	108 (0.6)	99 (0.6)	28 (1.5)	21 (1.1)	49 (2.5)	37 (1.9)

Memory loss, n (%)												
No	22,752 (99.8)	22,716 (99.7)	11,622 (99.9)	11,621 (99.9)	2,872 (99.8)	2,870 (99.8)	16,626 (99.8)	16,599 (99.7)	1,833 (99.8)	1,831 (99.7)	1,951 (99.5)	1,946 (99.3)
Yes	36 (0.2)	72 (0.3)	11 (0.1)	12 (0.1)	5 (0.2)	7 (0.2)	27 (0.2)	54 (0.3)	4 (0.2)	6 (0.3)	9 (0.5)	14 (0.7)
Confusion or disorientation, n (%)												
No	22,699 (99.6)	22,706 (99.6)	11,621 (99.9)	11,617 (99.9)	2,869 (99.7)	2,869 (99.7)	16,531 (99.3)	16,526 (99.2)	1,817 (98.9)	1,817 (98.9)	1,929 (98.4)	1,939 (98.9)
Yes	89 (0.4)	82 (0.4)	12 (0.1)	16 (0.1)	8 (0.3)	8 (0.3)	122 (0.7)	127 (0.8)	20 (1.1)	20 (1.1)	31 (1.6)	21 (1.1)
Dementia, n (%)												
No	22,694 (99.6)	22,709 (99.7)	11,628 (100.0)	11,625 (99.9)	2,870 (99.8)	2,872 (99.8)	16,494 (99.0)	16,494 (99.0)	1,810 (98.5)	1,816 (98.9)	1,944 (99.2)	1,947 (99.3)
Yes	94 (0.4)	79 (0.3)	5 (0.0)	8 (0.1)	7 (0.2)	5 (0.2)	159 (1.0)	159 (1.0)	27 (1.5)	21 (1.1)	16 (0.8)	13 (0.7)
Cardiac arrhythmia, n (%)												
No	22,627 (99.3)	22,598 (99.2)	11,594 (99.7)	11,593 (99.7)	2,860 (99.4)	2,850 (99.1)	16,515 (99.2)	16,488 (99.0)	1,819 (99.0)	1,816 (98.9)	1,935 (98.7)	1,931 (98.5)

Yes	161 (0.7)	190 (0.8)	39 (0.3)	40 (0.3)	17 (0.6)	27 (0.9)	138 (0.8)	165 (1.0)	18 (1.0)	21 (1.1)	25 (1.3)	29 (1.5)
Respiratory, n (%)												
No	20,942 (91.9)	20,942 (91.9)	11,113 (95.5)	11,149 (95.8)	2,555 (88.8)	2,566 (89.2)	14,529 (87.2)	15,054 (90.4)	1,541 (83.9)	1,615 (87.9)	1,413 (72.1)	1,579 (80.6)
Yes	1,846 (8.1)	1,846 (8.1)	520 (4.5)	484 (4.2)	322 (11.2)	311 (10.8)	2,124 (12.8)	1,599 (9.6)	296 (16.1)	222 (12.1)	547 (27.9)	381 (19.4)
CV, n (%)												
No	22,349 (98.1)	22,313 (97.9)	11,522 (99.3)	11,541 (99.2)	2,823 (98.1)	2,820 (98.0)	16,068 (96.5)	16,035 (96.3)	1,735 (94.4)	1,738 (94.6)	1,832 (93.5)	1,828 (93.3)
Yes	439 (1.9)	475 (2.1)	81 (0.7)	92 (0.8)	54 (1.9)	57 (2.0)	585 (3.5)	618 (3.7)	102 (5.6)	99 (5.4)	128 (6.5)	132 (6.7)
Mental health, n (%)												
No	21,848 (95.9)	21,707 (95.3)	11,402 (98.0)	11,340 (97.5)	2,733 (95.0)	2,723 (94.6)	15,834 (95.1)	15,787 (94.8)	1,718 (93.5)	1,719 (93.6)	1,793 (91.5)	1,816 (92.7)
Yes	940 (4.1)	1,081 (4.7)	231 (2.0)	293 (2.5)	144 (5.0)	154 (5.4)	819 (4.9)	866 (5.2)	119 (6.5)	118 (6.4)	167 (8.5)	144 (7.3)
Cancer, n (%)												

No	22,561 (99.4)	22,637 (99.3)	11,609 (99.8)	11,603 (99.7)	2,859 (99.4)	2,863 (99.5)	16,558 (99.4)	16,547 (99.4)	1,828 (99.5)	1,825 (99.3)	1,938 (98.9)	1,946 (99.3)
Yes	137 (0.6)	151 (0.7)	24 (0.2)	30 (0.3)	18 (0.6)	14 (0.5)	95 (0.6)	106 (0.6)	9 (0.5)	12 (0.7)	22 (1.1)	14 (0.7)
Respiratory and CV, n (%)												
No	22,657 (99.4)	22,664 (99.5)	11,617 (99.9)	11,603 (99.7)	2,860 (99.4)	2,858 (99.3)	16,452 (98.8)	16,472 (98.9)	1,807 (98.8)	1,811 (98.6)	1,897 (96.8)	1,913 (97.6)
Yes	131 (0.6)	124 (0.5)	16 (0.1)	30 (0.3)	17 (0.6)	19 (0.7)	201 (1.2)	181 (1.1)	30 (1.6)	26 (1.4)	63 (3.2)	47 (2.4)
Respiratory and mental health, n (%)												
No	22,583 (99.1)	22,572 (99.1)	11,568 (99.4)	11,570 (99.5)	2,829 (98.3)	2,844 (98.9)	16,391 (98.4)	16,408 (98.5)	1,792 (97.6)	1,803 (98.1)	1,877 (95.8)	1,910 (97.4)
Yes	205 (0.9)	216 (0.9)	65 (0.6)	63 (0.5)	48 (1.7)	33 (1.1)	262 (1.6)	245 (1.5)	45 (2.4)	34 (1.9)	83 (4.2)	50 (2.6)
Mental health and CV, n (%)												
No	22,735 (99.8)	22,739 (99.8)	11,625 (99.9)	11,624 (99.9)	2,873 (99.9)	2,875 (99.9)	16,596 (99.7)	16,589 (99.6)	1,827 (99.5)	1,826 (99.4)	1,953 (99.6)	1,949 (99.4)
Yes	53 (0.2)	49 (0.2)	8 (0.1)	9 (0.1)	4 (0.1)	2 (0.1)	57 (0.3)	64 (0.4)	10 (0.5)	11 (0.6)	7 (0.4)	11 (0.6)

Respiratory, CV, and mental health, n (%)												
No	22,731 (99.7)	22,736 (99.8)	11,624 (99.9)	11,623 (99.9)	2,871 (99.8)	2,868 (99.7)	16,555 (99.4)	16,569 (99.5)	1,820 (99.1)	1,822 (99.2)	1,930 (98.5)	1,929 (98.4)
Yes	57 (0.3)	52 (0.2)	9 (0.1)	10 (0.1)	6 (0.2)	9 (0.3)	98 (0.6)	84 (0.5)	17 (0.9)	15 (0.8)	30 (1.5)	31 (1.6)
No conditions, n (%)												
No	2,722 (11.9)	2,909 (12.8)	725 (6.2)	747 (6.4)	439 (15.3)	450 (15.6)	2,812 (16.9)	2,425 (14.6)	398 (21.7)	338 (18.4)	629 (32.1)	487 (24.8)
Yes	20,066 (88.1)	19,879 (87.2)	10,908 (93.8)	10,886 (93.6)	2,438 (84.7)	2,427 (84.4)	13,841 (83.1)	14,228 (85.4)	1,439 (78.3)	1,499 (81.6)	1,331 (67.9)	1,473 (75.2)

COPD, chronic obstructive pulmonary disease; COVID-19, coronavirus disease 2019; CV, cardiovascular; ER, emergency room; ICU, intensive care unit

Supplemental Table 5 Full list of risk ratios and 95% confidence intervals of covariates associated with the occurrence of new respiratory, cardiovascular, and mental health conditions at >30–≤90 days or >90–≤180 days post COVID-19 diagnosis or hospital discharge

RR (95% CI)	Respiratory conditions		Cardiovascular conditions		Mental health conditions	
	>30–≤90 days	>90–≤180 days	>30–≤90 days	>90–≤180 days	>30–≤90 days	>90–≤180 days
Age group						
18–29 years	1.18 (0.80, 1.79)	1.08 (0.76, 1.56)	NA*	3.18 (0.94, 19.53)	2.67 (1.59, 4.85)	1.81 (1.22, 2.79)
30–39 years	1.90 (1.32, 2.82)	1.59 (1.14, 2.27)	NA*	6.09 (1.88, 36.34)	3.00 (1.79, 5.43)	1.87 (1.26, 2.88)
40–49 years	2.44 (1.71, 3.58)	1.96 (1.41, 2.77)	NA*	6.68 (2.05, 39.85)	2.73 (1.63, 4.96)	1.91 (1.29, 2.94)
50–64 years	2.84 (2.01, 4.13)	1.92 (1.40, 2.71)	NA*	8.03 (2.52, 47.28)	2.94 (1.77, 5.28)	1.63 (1.11, 2.50)
65–74 years	2.68 (1.86, 3.95)	1.88 (1.34, 2.70)	NA*	10.46 (3.10, 62.22)	2.30 (1.35, 4.22)	1.28 (0.85, 2.02)
75–84 years	2.43 (1.65, 3.65)	1.52 (1.04, 2.24)	NA*	18.40 (5.23, 107.58)	2.61 (1.50, 4.85)	1.29 (0.83, 2.07)
≥85 years	1.80 (1.10, 2.95)	1.62 (1.01, 2.58)	NA*	15.49 (2.47, 111.25)	3.05 (1.59, 6.02)	1.73 (1.00, 2.98)
Sex						
Male	0.91 (0.84, 0.99)	0.84 (0.77, 0.92)	1.18 (0.92, 1.51)	1.26 (0.98, 1.61)	0.63 (0.56, 0.70)	0.53 (0.47, 0.59)
Race						

Caucasian	1.02 (0.94, 1.12)	1.13 (1.02, 1.24)	1.10 (0.83, 1.47)	1.10 (0.83, 1.48)	1.48 (1.31, 1.69)	1.792 (1.59, 2.03)
Asian	0.99 (0.81, 1.19)	1.00 (0.80, 1.24)	0.72 (0.33, 1.36)	0.84 (0.41, 1.56)	0.74 (0.50, 1.04)	0.81 (0.56, 1.12)
Ethnicity						
Non-Hispanic	1.12 (0.97, 1.29)	1.38 (1.17, 1.62)	2.56 (1.53, 4.63)	2.12 (1.30, 3.70)	1.40 (1.14, 1.73)	1.54 (1.27, 1.87)
Obesity	1.33 (1.21, 1.46)	1.38 (1.24, 1.53)	1.08 (0.76, 1.51)	1.18 (0.83, 1.63)	1.41 (1.24, 1.59)	1.49 (1.32, 1.67)
Insurance						
Medicaid	1.01 (0.89, 1.15)	1.07 (0.94, 1.22)	1.24 (0.85, 1.76)	1.23 (0.84, 1.75)	1.39 (1.18, 1.64)	1.41 (1.21, 1.63)
Medicare	1.06 (0.93, 1.21)	1.06 (0.91, 1.22)	1.45 (0.94, 2.21)	1.51 (1.00, 2.26)	1.53 (1.29, 1.82)	1.46 (1.23, 1.72)
Other payor type	0.77 (0.64, 0.93)	1.07 (0.90, 1.27)	1.00 (0.56, 1.65)	1.10 (0.64, 1.77)	0.97 (0.75, 1.23)	1.03 (0.82, 1.28)
Uninsured	0.71 (0.58, 0.85)	0.65 (0.52, 0.80)	0.96 (0.55, 1.57)	0.66 (0.33, 1.16)	0.76 (0.56, 0.99)	0.63 (0.47, 0.82)
Sub-cohort						
ER on diagnosis	0.64 (0.56, 0.74)	0.56 (0.48, 0.65)	0.45 (0.27, 0.71)	0.59 (0.38, 0.89)	0.60 (0.49, 0.72)	0.65 (0.55, 0.78)
ER	1.39 (1.17, 1.65)	1.33 (1.10, 1.58)	1.41 (0.83, 2.27)	1.57 (0.95, 2.47)	1.22 (0.95, 1.54)	1.20 (0.95, 1.50)
Hospitalisation without ICU	1.33 (1.20, 1.47)	1.02 (0.91, 1.14)	1.74 (1.28, 2.35)	1.42 (1.04, 1.94)	1.03 (0.89, 1.18)	1.08 (0.95, 1.23)
ICU without ventilation	1.69 (1.39, 2.03)	1.18 (0.93, 1.47)	1.69 (0.75, 3.28)	2.41 (1.25, 4.23)	1.31 (0.99, 1.71)	1.34 (1.02, 1.73)

ICU with ventilation	2.64 (2.27, 3.04)	1.86 (1.55, 2.21)	3.16 (1.83, 5.18)	2.65 (1.49, 4.43)	1.89 (1.51, 2.35)	1.52 (1.20, 1.91)
Month of COVID-19 diagnosis						
Feb–Apr 2020	1.08 (0.97, 1.20)	1.06 (0.93, 1.19)	0.88 (0.62, 1.24)	0.96 (0.66, 1.38)	0.85 (0.72, 1.01)	0.93 (0.80, 1.09)
May 2020	0.75 (0.67, 0.83)	0.92 (0.82, 1.04)	0.76 (0.55, 1.06)	1.21 (0.89, 1.65)	0.82 (0.71, 0.95)	0.90 (0.79, 1.03)
Jun 2020	0.58 (0.51, 0.65)	0.72 (0.63, 0.81)	0.68 (0.48, 0.96)	0.75 (0.52, 1.07)	0.72 (0.62, 0.84)	0.74 (0.64, 0.86)
Jul 2020	0.48 (0.34, 0.65)	0.61 (0.45, 0.82)	0.33 (0.08, 0.88)	0.81 (0.34, 1.65)	0.60 (0.40, 0.87)	0.79 (0.57, 1.08)
Weighted CCI	1.05 (1.02, 1.07)	1.07 (1.04, 1.09)	1.17 (1.10, 1.25)	1.16 (1.08, 1.23)	1.12 (1.09, 1.15)	1.14 (1.11, 1.17)

Grey and blue shading denote increased and decreased risk of a new condition occurring, respectively. Reference groups are: <18 years (age), female (sex), African American (race), Hispanic (ethnicity), non-obese (obesity), diagnosis in February and March 2020 (diagnosis month); commercial (insurance), outpatient (sub-cohort).

*Values were not calculable as the reference group (<18 years) had no new diagnoses of clinical conditions

CCI, Charlson Comorbidity Index; CI, confidence interval; COVID-19, coronavirus disease 2019; ER, emergency room; ICU, intensive care unit;

NA, not applicable; RR, risk ratio

Supplemental Table 6 Full list of risk ratios and 95% confidence intervals of covariates associated with a new cancer diagnosis at >30–≤90 days or >90–≤180 days post COVID-19 diagnosis or hospital discharge

	Risk ratio (95% CI)	
	>30–≤90 days	**>90–≤180 days**
Age group		
18–29 years	0.95 (0.31, 4.16)	1.17 (0.19, 22.13)
30–39 years	0.89 (0.29, 3.90)	2.63 (0.51, 46.94)
40–49 years	1.43 (0.49, 6.05)	3.38 (0.69, 59.46)
50–64 years	2.44 (0.90, 9.93)	8.35 (1.84, 138.73)
65–74 years	3.19 (1.13, 13.21)	13.50 (2.91, 217.30)
75–84 years	2.60 (0.86, 11.12)	15.50 (3.25, 247.65)
≥85 years	2.12 (0.53, 10.33)	8.45 (1.39, 151.23)
Sex		
Male	0.97 (0.75, 1.27)	1.03 (0.79, 1.33)
Race		
Caucasian	1.51 (1.10, 2.11)	1.05 (0.78, 1.43)
Asian	1.45 (0.66, 2.80)	1.15 (0.53, 2.20)
Ethnicity		
Non-Hispanic	2.49 (1.33, 5.28)	1.34 (0.79, 2.50)

Obesity	1.34 (0.98, 1.80)	1.34 (1.00, 1.79)
Insurance		
Medicaid	0.88 (0.52, 1.41)	0.68 (0.38, 1.13)
Medicare	1.43 (0.99, 2.07)	1.10 (0.78, 1.55)
Other payor type	0.72 (0.32, 1.39)	0.69 (0.32, 1.29)
Uninsured	1.32 (0.68, 2.32)	0.49 (0.17, 1.09)
Sub-cohort		
ER on diagnosis	0.47 (0.27, 0.76)	0.53 (0.32, 0.83)
ER	1.23 (0.68, 2.06)	0.69 (0.32, 1.30)
Hospitalization without ICU	0.71 (0.50, 0.99)	0.65 (0.47, 0.90)
ICU without ventilation	0.45 (0.16, 1.01)	0.50 (0.21, 1.02)
ICU with ventilation	1.24 (0.71, 2.06)	0.56 (0.27, 1.04)
Month of COVID-19 diagnosis		
Feb–Apr 2020	1.44 (0.98, 2.11)	1.72 (1.17, 2.51)
May 2020	1.21 (0.85, 1.72)	1.18 (0.82, 1.69)
Jun 2020	0.90 (0.61, 1.32)	1.28 (0.89, 1.84)
Jul 2020	0.70 (0.21, 1.71)	0.80 (0.24, 1.96)
Weighted CCI	1.04 (0.97, 1.11)	1.06 (0.99, 1.12)

Reference groups are: <18 years (age), female (sex), African American (race), Hispanic (ethnicity), non-obese (obesity); diagnosis in February and March 2020 (diagnosis month), commercial (insurance), outpatient (sub-cohort).

CCI, Charlson Comorbidity Index; CI, confidence interval;
COVID-19, coronavirus disease 2019; ER, emergency room; ICU, intensive care unit

The RECORD statement – checklist of items, extended from the STROBE statement, that should be reported in observational studies using routinely collected health data.

	Item No.	STROBE items	Location in manuscript where items are reported	RECORD items	Location in manuscript where items are reported
Title and abstract					
	1	(a) Indicate the study's design with a commonly used term in the title or the abstract (b) Provide in the abstract an informative and balanced summary of what was done and what was found	*(a) confirmed (Design section) (b) confirmed adequately covered in abstract*	RECORD 1.1: The type of data used should be specified in the title or abstract. When possible, the name of the databases used should be included. RECORD 1.2: If applicable, the geographic region and timeframe within which the study took place should be reported in the title or abstract. RECORD 1.3: If linkage between databases was conducted for the study, this should be clearly stated in the title or abstract.	*Abstract (Objective and Design)* *Abstract (Setting and Participants)* *Not applicable*
Introduction					
Background rationale	2	Explain the scientific background and rationale for the investigation being reported	*Confirmed present in introduction*		
Objectives	3	State specific objectives, including any prespecified hypotheses	*Specific objective stated (last paragraph of introduction; there were no pre-*		

			specified hypotheses)		
Methods					
Study Design	4	Present key elements of study design early in the paper	*Included in methods ('Patients and study design') and described in Figure 1*		
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	*Included in methods ('Database' & 'Patients and study design' sections)*		

Participants	6	*(a) Cohort study*- Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*- Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*- Give the eligibility criteria, and the sources and methods of selection of participants *(b) Cohort study*- For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*- For matched studies, give matching criteria and the number of controls per case	*(a) confirmed included in methods ('Patients and study design' section (b) not relevant (not a matched study)*	RECORD 6.1: The methods of study population selection (such as codes or algorithms used to identify subjects) should be listed in detail. If this is not possible, an explanation should be provided. RECORD 6.2: Any validation studies of the codes or algorithms used to select the population should be referenced. If validation was conducted for this study and not published elsewhere, detailed methods and results should be provided. RECORD 6.3: If the study involved linkage of databases, consider use of a flow diagram or other graphical display to demonstrate the data linkage process, including the number of individuals with linked data at each stage.	*Confirmed in methods ('Patients and study design')* *The algorithms have been used previously and is cited in the methods (Chawla et al., 2021)* *Not applicable*
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable.	*All definitions are presented in the methods ('Patients and study design', 'Modelling and*	RECORD 7.1: A complete list of codes and algorithms used to classify exposures, outcomes, confounders, and effect modifiers should be provided. If these cannot be reported, an explanation should be provided.	*Relevant lists are provided throughout the manuscript (e.g. ICD-10 codes in*

			statistical analysis', and 'Sensitivity analysis' sections)		*supplemental Tables 1 and 2, list of confounders in methods section 'Modeling and statistical analysis')*
Data sources/ measurement	8	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	*Source of data is the Optum Electronic Medical Record data, and are routinely collected by practicing physicians (detailed in methods section)*		
Bias	9	Describe any efforts to address potential sources of bias	*A sensitivity analysis was performed, and relevant controls (non-hospital setting covariates) were included in our statistical models*		
Study size	10	Explain how the study size was arrived at	*All eligible patients in the Optum dataset were included, without a prespecified study size (explained in the*		

			database and patients and study design section)		
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why	*Described in Methods section 'Modeling and statistical analysis'*		
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding (b) Describe any methods used to examine subgroups and interactions (c) Explain how missing data were addressed (d) *Cohort study*- If applicable, explain how loss to follow-up was addressed *Case-control study*- If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*- If applicable, describe analytical methods taking account of sampling strategy (e) Describe any sensitivity analyses	*a) Methods ('Modeling and statistical analysis')* *b) We do not conduct sub-group analysis* *c) Explained in discussion section* *d) We have conducted a retrospective cohort study. Regarding the loss-to follow-up, since we are not assessing the effect of a treatment, rather looking at disease severity, we assume it is non-differential.* *e) as described in methods ('Sensitivity analysis')*		

Data access and cleaning methods		..		RECORD 12.1: Authors should describe the extent to which the investigators had access to the database population used to create the study population.	*Authors had access to deidentified EMR data*
				RECORD 12.2: Authors should provide information on the data cleaning methods used in the study.	*Data cleaning methods have been described previously; the reference is cited in the Methods 'Database' section (Chawla et al).*
Linkage		..		RECORD 12.3: State whether the study included person-level, institutional-level, or other data linkage across two or more databases. The methods of linkage and methods of linkage quality evaluation should be provided.	*EMR data from hospital networks were used to form the Optum dataset. Linkage of EMR data and methods are described on Optum's website: https://www.optum.com/business/solutions/life-sciences/real-world-data/ehr-data.html*

Results					
Participants	13	(a) Report the numbers of individuals at each stage of the study (*e.g.*, numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed) (b) Give reasons for non-participation at each stage. (c) Consider use of a flow diagram	*The number of patients in the dataset and those with a COVID-19 diagnosis is given in the Methods 'Database' section. The criteria on how the population is selected is made clear in methods 'Patients and study design' section (inclusion and exclusion criteria).*	RECORD 13.1: Describe in detail the selection of the persons included in the study (*i.e.,* study population selection) including filtering based on data quality, data availability and linkage. The selection of included persons can be described in the text and/or by means of the study flow diagram.	*The criteria on how the population is selected is made clear in methods 'Patients and study design' section (inclusion and exclusion criteria)*
Descriptive data	14	(a) Give characteristics of study participants (*e.g.*, demographic, clinical, social) and information on exposures and potential confounders (b) Indicate the number of participants with missing data for each variable of interest (c) *Cohort study*- summarise follow-up time (*e.g.*, average and total amount)	*Shown in detail in tables e.g. Table 1 and also in appendix (table 1 and 2)*		
Outcome data	15	*Cohort study*- Report numbers of outcome events or summary measures over time *Case-control study*-	*Outcome data are presented in Table 2*		

		Report numbers in each exposure			
		category, or summary measures of exposure *Cross-sectional study*- Report numbers of outcome events or summary measures			
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder- adjusted estimates and their precision (e.g., 95% confidence interval). Make clear which confounders were adjusted for and why they were included (b) Report category boundaries when continuous variables were categorized (c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	*Confounder-adjusted estimates are provided; however, unadjusted results could be derived from Table 2, where the raw counts and percentages can be used to calculate raw measures of effect. Confounders we control for are described in the modelling and statistical analysis section*		
Other analyses	17	Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses	*Sensitivity analysis is reported*		

Discussion					
Key results	18	Summarise key results with reference to study objectives	*Covered in discussion*		
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	*An extensive limitations section is included, covering the relevant aspects*	RECORD 19.1: Discuss the implications of using data that were not created or collected to answer the specific research question(s). Include discussion of misclassification bias, unmeasured confounding, missing data, and changing eligibility over time, as they pertain to the study being reported.	*An extensive limitations section is included, covering the relevant aspects*
Interpretation	20	Give a cautious overall interpretation of results considering objectives,	*Covered in discussion*		
		limitations, multiplicity of analyses, results from similar studies, and other relevant evidence			
Generalisability	21	Discuss the generalisability (external validity) of the study results	*Covered in discussion*		
Other Information					
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present	*Covered in funding section*		

		article is based			
Accessibility of protocol, raw data, and programming code		..		RECORD 22.1: Authors should provide information on how to access any supplemental information such as the study protocol, raw data, or programming code.	*Information is included in the data availability statement*

*Reference: Benchimol EI, Smeeth L, Guttmann A, Harron K, Moher D, Petersen I, Sørensen HT, von Elm E, Langan SM, the RECORD Working Committee. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. *PLoS Medicine* 2015; in press.